

TEMPERAMENTAL DIFFERENCES IN MAZE PERFORMANCE

PART I. CHARACTERISTIC DIFFERENCES AMONG PSYCHONEUROTICS

By G. A. FOULDS, *Runwell Hospital*

I. Introduction (p. 209). II. Modifications in the administration of the Porteus Mazes (pp. 209-211). III. Modifications in the scoring (p. 211). IV. Differences in performance due to sex, age and psychoneurosis (pp. 211-214). V. Comparison of the results of the psychoneurotic subgroups (pp. 214-216). VI. Summary and outlook (pp. 216-217). References (p. 217).

I. INTRODUCTION

Psychiatrists frequently look to clinical psychologists to provide some objective measurement of the depth or intensity of affective disturbances such as anxiety and depression. The present investigation, reported in this and a subsequent paper, attempts partially to meet this need and to help clarify the problem of psychomotor retardation.

Objective measurements can, of course, be made only of the presumed effects of a subjective state. The most relevant subjects among psychoneurotics are those commonly diagnosed as anxiety states or reactive depressions (Eysenck's⁽⁴⁾ 'dysthymics', 1947). Some relevant symptoms appear to be a reduction in the ability (i) to arrive at a firm decision (vacillation); (ii) to think or act at a speed sufficient to allow of adjustment to a given situation (retardation); (iii) to maintain adequate muscular control and co-ordination (tremor); (iv) to concentrate; (v) to recall recently acquired information.

The Porteus Maze Test, with some modifications in the administration and scoring, appeared likely to provide the possibility of measuring the effects of these symptoms. If such measurements could be made, the test should theoretically differentiate between dysthymics on the one hand and hysterics and psychopaths on the other. This, therefore, constitutes the first stage in the attempted validation of the test procedure.

Since affective disturbance implies some degree of awareness of one's own mental state, the attention of dysthymics should be divided more frequently than that of hysterics or psychopaths. This would appear to be a sufficient explanation for some of the symptoms mentioned above, notably vacillation, retardation and lack of concentration. The assumption is made that a quasi-automatic activity, when combined with performance of the test, would serve temporarily to obliterate awareness of the affective disturbance, and that consequently the performance of the dysthymics would become much more like that of the hysterics and psychopaths. This constitutes the second stage in the attempted validation of the test procedure. A third stage could be provided by studying changes in test performance concomitant with changes in the clinical condition of the subjects.

II. MODIFICATIONS IN THE ADMINISTRATION OF THE PORTEUS MAZES

Recent developments in the use of the Porteus Mazes (Porteus⁽⁶⁾, 1942) and Le Test du Labyrinthe (Chapuis⁽²⁾, 1949) have emphasized the manner in which these tests are done, whilst still recognizing the value of assessing the ability to do them.

Chapuis⁽²⁾ regards his test as a test of intelligence and character: 'Une épreuve, d'intelligence il peut s'agir aussi d'un problème de mémoire, d'adresse manuelle, etc. n'est point, et cela même, un bon test de caractère. Tout dépendra, d'une part, des possibilités d'observation qu'elle offre à l'examineur et, d'autre part, de l'influence réelle des facteurs

caractériels sur le rendement obtenu' (p. 66). Consequently, he uses the speed/accuracy ratio to give a global score. For the remaining character traits he is content to rely on subjective evaluation, when, in many cases, objective measurements could have been made without distortion of the findings. The italicized part of the above statement (presented in the writer's italics) leads to serious confusion. Any influence of cognitive on orectic functions, or the converse, should be regarded as a limitation rather than as an additional merit in a test.

In the Porteus Mazes deductions are made from the number of times the subject lifts his pencil, cuts corners, crosses lines, etc., contrary to instructions. Such failure to comply with the prescribed manner of doing the test can be scored objectively and correlated with the ability to thread the mazes correctly. It would, therefore, appear legitimate to consider what Porteus calls the qualitative and quantitative aspects separately.

The ability to thread mazes is thought by Porteus to depend on qualities that are best described by such terms as prudence, forethought, planning rather than intelligence. He points out that many psychologists have tended to regard the Maze Test as a test of some aspect of intelligence rather than as a temperament test, and this he attributes to the fact that in the majority of investigations the test has correlated as highly as +0.6 to +0.7 with the Binet Test (Porteus(7), 1946)¹. In addition, he states that its interpretation as a test of temperament would be aided by the further separation of the intellectual from the temperamental factors involved (Porteus(6), 1942).

The modifications to be reported in this paper, which have been made in the administration and scoring of this test have had as their principal object the heightening of the temperamental and the lowering of the intellectual factors. Porteus warns that timing the test alters its nature (Porteus(7), 1946). For the reason mentioned above, this, therefore, constitutes the main modification in the administration. Times were recorded, to the nearest second, from the instruction to start to the commencement of tracing (starting time) and from the beginning to the end of tracing (tracing time). The total tracing time was recorded in units of 15 sec. In Babcock's(4) view 'there is a causal relation between time as shown in poor efficiency' (on a variety of tests) 'and lack of ability to make normal adjustments' (Babcock(4), 1941, p. 146). Even those who do not accept all the implications of this statement will agree that an unduly high proportion of psychoneurotic and psychotic subjects deviate markedly from the normal tempo in many diverse activities. It was, therefore, anticipated that these two scoring categories would be among the most important.

As a necessary consequence of timing the test, the same maze form is used throughout and is not removed after a mistake has been made. A pilot survey soon demonstrated that some subjects made so many mistakes that the record was rendered indecipherable. Subjects were, therefore, stopped after two wrong directions had been taken in years 5-11 and after four in years 12 and 14 and adult I and II. Since the correlation between total wrong direction score and total time taken was as low as +0.27 with sixty normal subjects and +0.02 with sixty psychoneurotics, this did not seriously interfere with the usefulness of timing and probably less so than if some subjects were permitted to penetrate every nook and cranny.

¹ In his most recent publication (Porteus(8), 1950) he defines intelligence as 'the capacity for making planned responses to an increasing range of relevant stimuli'. He has no longer any hesitation in believing that a test which measures something different from the Binet tests can still be a test of intelligence and states that he now considers the Maze Test to be a test of intelligence.

The complete instructions given to subjects were as follows:

Year 5 presented. Tester says: In this test you have to start at the letter *S* and trace your way through the maze and out at the one opening provided. *You must not lift your pencil, cross any lines, nor go up any road that is blocked.* Should you do so by mistake, you still must not lift your pencil, but should trace your way back again. You must not trace the maze in the air. You may only use one hand in working at the problem; the other should be used to steady the paper. In each case wait for the word 'Go'. *Right, start at the S and trace your way through. Go.*

Years 6 and 7. Repeat the italicized phrases above.

Years 8, etc. Say simply: Are you ready? Go.

The Tester demonstrates with year 5 by tracing, with great care, a course down the middle of the first passage and at right angles round the first bend.

III. MODIFICATIONS IN THE SCORING

It was considered that it would be of interest to assess not only the number of times that subject embarked on a wrong passage, but how far he went before he discovered his error. All wrong passages were, therefore, marked off on a master-key form into three sections, and a weighting was given in accordance with the number of sections entered. The writer subsequently discovered that he had been anticipated in this device by Chapuis (2949). Chapuis does not, however, consider the ratio between the wrong direction score and the weighted wrong direction score.

Porteus weights lifted pencils, cut corners, etc., on the basis of differences found between normal and delinquent children, and then sums the scores from each category to give a global qualitative score. This seems to the present writer to be a mating of incompatibles and psychologically meaningless, particularly as the different scores are given an additional weighting by the varied means and standard deviations. These categories are, therefore, in the present instance kept separate, with the exception of cut corners and crossed lines, which correlated +0.74 and followed the same pattern in all the groups tested.

In accordance with the practice of Porteus, lifted pencils or crossed lines in the first third of each test were noted separately. (Crossed lines were subsequently taken to include cut corners and crossed lines.)

IV. DIFFERENCES IN PERFORMANCE DUE TO SEX, AGE AND PSYCHONEUROSIS

Sex

The male group comprised twenty-two members of the Research Unit of a photographic manufacturing company and eighteen psychiatric nurses; the female group consisted of nine from the Research Unit, thirteen psychiatric nurses and eighteen occupational therapists. Since the similarities and differences between occupational therapists and male research workers are paralleled by those between male and female nurses, the sample is probably valid for the present purpose. The groups were matched for age and Progressive Matrices score (1938, untimed): males 24.25 years ± 4.52 and 49.92 ± 8.12 ; females 24.30 years ± 4.79 and 50.25 ± 7.52 .

The following abbreviations are used throughout the tables:

St.T.	= starting time.	L.P.	= lifted pencils.
T.T.	= tracing time.	C.L.	= crossed lines.
W.D.	= wrong direction taken.	W.L.	= wavy lines.
W.W.D.	= weighted wrong direction score.		

Table 1. *Maze results of forty male and forty female subjects*

		St.T.	T.T.	W.D. in 1st third of any maze	W.D.	W.W.D.	$\frac{W.D.}{W.W.D.} \times \frac{100}{1}$	L.P.	C.L.	W.L.
40 men:	Mean	22.50	19.90	35%	9.06	16.22	55.9	4.77	13.45	2.82
	S.D.	11.51	6.90	—	5.39	9.99	—	4.74	10.97	1.88
40 women:	Mean	26.62	25.00	36%	10.32	16.90	61.1	5.67	12.02	2.47
	S.D.	14.44	11.96	—	4.97	9.79	—	4.23	10.35	2.52

The only statistically significant difference between means was in total tracing time, the male group working more quickly (C.R. 2.34). The difference between the mean starting times was 1.41 times the standard error of the difference. In the remaining categories this value was still lower.

Twelve men as against three women lifted their pencils or crossed lines in the first two mazes (for $n-1$, $\chi^2=6.65$, a highly significant difference)¹. Six men as against one woman had a weighted wrong direction (W.W.D.) score of more than twice the wrong direction (W.D.) score ($\chi^2=3.91$). There were, therefore, more men who made a large number of gross errors. If we consider impulsiveness to be responsible for a loss of accuracy associated with speed, then more young men than young women showed themselves to be impulsive.

Age

A group with a mean age of 33.30 years (S.D. 6.40) was compared with a group having a mean age of 20.42 (S.D. 2.25). The respective Progressive Matrices scores were 50.97 (S.D. 7.83) and 50.57 (S.D. 6.22). The older group comprised twenty-two members of the photographic unit and eighteen hospital staff, fourteen being women; in the younger group the corresponding numbers were twenty-one, nineteen and eighteen. Maze scores are shown in Table 2.

Table 2. *Maze results of forty subjects over 25 years and forty subjects 25 years and under*

		St.T.	T.T.	W.D. in 1st third of any maze	W.D.	W.W.D.	$\frac{W.D.}{W.W.D.} \times \frac{100}{1}$	L.P.	C.L.	W.L.
40 older:	Mean	35.37	21.47	35%	7.77	12.50	62.2	4.55	10.52	2.42
	S.D.	19.65	8.41	—	5.04	9.92	—	4.01	8.75	2.29
40 younger:	Mean	19.95	21.47	35%	10.65	18.22	58.5	5.72	12.95	2.37
	S.D.	10.39	10.71	—	4.61	8.67	—	5.17	8.25	1.85

The older subjects spent a significantly longer time considering the problem before starting to trace (C.R. 4.39). They then proceeded to work just as quickly as the young subjects. Their initial prudence, however, resulted in fewer wrong directions being taken (C.R. 2.66). In addition, there were significantly more of the younger group who made gross errors, 7 to 1 ($\chi^2=5.00$). Twelve of the younger group against three of the older lifted pencils or crossed lines in either or both of the first two mazes ($\chi^2=6.65$).

¹ All subsequent χ^2 values are for one degree of freedom. Values greater than 3.84 are, therefore, significant at least at the 5% level of confidence.

Psychoneurosis

Sixty normal and sixty psychoneurotic subjects were next compared. Again the groups were matched for age and Progressive Matrices score. The mean age of the normal group was 27.35 (s.d. 8.31) and of the psychoneurotics 27.35 (s.d. 7.42). The mean Matrices scores were 44.48 (s.d. 9.55) and 44.60 (s.d. 9.90), an approximately normal Matrices distribution. The normal group contained thirty-one and the psychoneurotic group thirty-two women. The latter group was unselected as regards nosological category, but, in fact, comprised twenty hysterics, twenty-two anxiety states, eleven reactive depressions and eleven obsessionals. This is a larger percentage of obsessionals than one would expect to get in a larger unselected group. The maze results are presented in Table 3.

Table 3. *Maze results of sixty normal and sixty psychoneurotic subjects*

	St.T.	T.T.	W.D. in 1st third of any maze	W.D.	W.W.D.	$\frac{W.D.}{W.W.D.} \times \frac{100}{1}$	L.P.	C.L.	W.L.
60 normals: Mean	29.78	24.02	34%	11.02	19.17	57.5	6.62	12.38	2.67
s.d.	21.77	10.71	—	5.16	10.93	—	5.13	10.06	2.07
60 neurotics: Mean	64.50	35.77	39%	9.65	17.63	54.7	9.30	12.85	3.28
s.d.	70.57	18.70	—	5.87	12.91	—	8.84	11.59	2.67

Normal subjects were significantly quicker both in starting and tracing (c.r. 3.64 and 2.23), and they tended to lift their pencils less often (c.r. 2.03); but the other scoring categories showed no significant differences between the groups.

Nine psychoneurotic subjects to two normals made more than half their errors of direction in the first third of any maze ($\chi^2=4.90$). Twenty psychoneurotics to ten normals gave a W.W.D. score of more than twice their W.D. score, i.e. have a large number of gross errors ($\chi^2=4.44$). Twenty-four psychoneurotics to twelve normals lifted their pencils over crossed lines in one or both of the first two mazes ($\chi^2=5.71$).

A comparison between the variances on the maze results for the two groups indicates greater scatter amongst the psychoneurotics. It may, therefore, be that the test is differentiating between types of temperament which are represented in their more extreme forms in the different psychoneurotic groups. The correlations presented in Table 4 between Progressive Matrices and the various maze scores at least suggest that temperamental factors enter more into psychoneurotic performance at the expense of the general intellectual factor. If this assumption be correct, the test may well be expected to differentiate more adequately between the psychoneurotic subgroups than between psychoneurotics and normals.

Table 4. *Correlations between Progressive Matrices and Porteus Mazes*

	St.T.	T.T.	W.D.	W.W.D.	$\frac{W.D.}{W.W.D.} \times \frac{100}{1}$	L.P.	C.L.	W.L.
60 normals: Matrices score	+0.28	-0.31	-0.64	-0.71	+0.46	-0.12	-0.08	0.00
60 neurotics: Matrices score	+0.14	-0.12	-0.55	-0.53	+0.29	+0.02	+0.02	+0.07

All correlations over 0.25 are significantly different from zero.

There is a slight tendency for those with high intellectual capacity to consider the mazes for a longer time before starting and then to work more quickly than those of lower capacity. They have a much more marked tendency to make fewer errors of direction and for these to be of a less gross type.

V. COMPARISON OF THE RESULTS OF THE PSYCHONEUROTIC SUBGROUPS

The relative scarcity of obsessional neurotics of low intellectual capacity made it impossible to match this group very closely with the other groups. The mean ages and Matrices scores were as follows: psychopaths, 35.47 (S.D. 6.52) and 38.63 (S.D. 11.69); hysterics, 35.20 (S.D. 9.92) and 38.38 (S.D. 10.60); anxiety states, 35.58 (S.D. 12.10) and 38.80 (S.D. 10.11); reactive depressions, 36.20 (S.D. 10.14) and 38.73 (S.D. 12.09); obsessionals, 35.92 (S.D. 11.44) and 43.56 (S.D. 8.71).

With the exception of the W.D. scores, the correlations with Progressive Matrices are not so high as to distort unduly the comparisons with the obsessional group. The group is, however, not comparable with the remaining groups in one other major particular. Sixteen of the twenty-five obsessionals were out-patients.¹ The subjects were much younger and worked very much more quickly than the hospital obsessionals. It may well be, however, that the groups are in fact more comparable on this account, since obsessionals are probably much more resistant to hospitalization than other psychoneurotics. While a random sample of in-patients will include relatively mild hysterics and anxiety states, it will include only severe obsessionals.

Table 5. *Scores of psychoneurotic subgroups on the Maze Test*

		St.T.	T.T.	W.D. in 1st third of any maze	W.D.	W.W.D.	$\frac{W.D.}{W.W.D.} \times \frac{100}{1}$	L.P.	C.L.	W.L.
30 psychopaths:										
	Mean	45.40	29.77	34.6%	12.03	20.87	57.6	7.00	11.80	2.77
	S.D.	32.30	12.77	—	5.52	11.77	—	6.41	8.49	1.77
45 hysterics:										
	Mean	44.22	34.49	36.3%	12.56	22.24	56.5	7.93	12.24	2.67
	S.D.	38.88	16.51	—	5.84	13.07	—	8.00	11.23	2.22
45 anxiety states:										
	Mean	76.71	37.27	38.4%	11.36	20.31	55.9	11.00	15.56	3.78
	S.D.	85.82	18.52	—	5.83	12.91	—	9.26	14.81	2.66
45 reactive depressions:										
	Mean	84.00	47.35	39.4%	10.27	18.73	54.8	11.49	8.29	3.09
	S.D.	88.55	25.53	—	6.10	13.88	—	11.98	7.92	2.68
25 obsessionals:										
	Mean	60.36	38.44	35.1%	8.44	14.24	59.3	5.96	6.88	1.44
	S.D.	57.47	15.75	—	3.84	7.56	—	4.77	6.51	1.60

Table 6 shows the critical ratios of the difference between the various means.

It will be seen that the test does not differentiate at all between psychopaths and hysterics. Whereas the sex distribution in the anxiety, depression and obsessional groups is much more even both here and in general, women predominate in hysteria and men in psychopathy. The largest of the differences obtained in the present experiment may well be due to this different sex-loading, since normal men were found to trace more quickly than normal women. It would appear, therefore, that there is no basic difference between these two

¹ The writer is greatly indebted to Dr Henry Wilson and Dr Davey of the London Hospital and to Dr No. Harris and Dr W. Haynes of the Middlesex Hospital for permission to test these patients.

Table 6. *Critical ratios of the difference between the means in Table 5*

	St.T.	T.T.	L.P.	C.L.	W.L.
Psychopaths					
Anxiety states	2.21	2.08	2.21	—	1.98
Psychopaths					
Depressions	2.67	3.94	2.10	(1.80)	—
Psychopaths					
Obsessionals	—	2.21	—	2.44	2.92
Hysterics					
Anxiety states	2.30	—	—	—	2.15
Hysterics					
Depressions	2.76	2.84	—	(1.94)	—
Hysterics					
Obsessionals	—	—	—	2.53	2.67
Anxiety states					
Depressions	—	2.14	—	2.91	—
Anxiety states					
Obsessionals	—	—	3.00	3.39	4.59
Depressions					
Obsessionals	—	—	2.72	—	3.22

groups in respect of the temperamental factors assessed by the Maze Test. On the Progressive Matrices and Thematic Apperception Test the tempo of hysterics and psychopaths shows a much greater degree of resemblance than does that of any other two groups. A similar result is obtained from the consideration of the ratio between Progressive Matrices and the Mill Hill vocabulary grade. These results, therefore, open up the possibility that such differences as exist between the two groups may, in part, be due to the cultural expectations of the individual and later of the psychiatrists making the diagnosis.

Mead (5) (1949), Dollard (3) and others (1944) have claimed that there are certain modes of behaviour which are considered by society to be appropriate to males and not to females and conversely. These modes of behaviour vary widely from one culture pattern to another. Adherence to these acceptable modes will normally be conducive to satisfactory personal and social adjustment. It is probable that there are certain deviations from the mode which are more acceptable than others. It may well be that hysteria, owing perhaps to the taboo on tenderness, is less acceptable in men in our society than is psychopathy. Certainly many psychopathic murderers, film stars and politicians are greatly admired by both sexes, albeit sometimes covertly, whereas male hysterics are considered unmanly and somewhat despicable. Female psychopaths, on the other hand, are probably less acceptable to our society than hysterics. They certainly were to the Victorians. They tend to be persecuted by other women, who, however, write about them and read about them in *Forever Amber*, *The Wicked Lady* and *Gone with the Wind*. To men they may sometimes be exciting, but never admirable.

To summarize the characteristic mazes of each of the nosological groups relative to the present test population:

*Psychopaths and hysterics.*¹ *Start without prolonged delay and then proceed rapidly with a relatively smooth, evenly flowing line. This is not, however, quite adequately under control since they cross a relatively large number of lines. They might be said to be somewhat impulsive and careless. Since the pencil is most frequently lifted at choice points and these subjects tend to hesitate less at such points, they have a low score for lifted pencils. On entering a cul-de-sac they characteristically wheel about immediately this is observed without manifest signs of tension, such as an increase in the waviness or pressure of the line or lifting of the pencil.*

Anxiety states. Perhaps because of their inability readily to make decisions, *start slowly and then work reasonably quickly, but their progress is jerky and accomplished only with a very tremulous line, which*

¹ Italics represent data scored on the test, the remainder being direct observations or inferences.

results in their crossing many lines. It is noticeable that the lines tend to become more wavy after an error has been committed. Brundage, in a private communication cited by Porteus (7) (1946), calls this type of reaction 'shatter'. Chapuis (2) (1949, p. 73) writes: 'L'anxiété nerveuse en présence de toute situation plus ou moins inédite se traduit fréquemment, au T.L., par un tracé peu sûr et tremblé ne s'expliquant point par des troubles graphiques proprement dits.'

Reactive depressions. The tempo is slow throughout. This perhaps enables them to avoid crossing an undue number of lines. Their defects appear to be defects of attention. They are slow to adjust to the total situation, making rather frequent errors in the early stages and forgetting to heed the instruction about lifting the pencil. There is a slight tendency for depressives to proceed further up blocked roads than other subjects. They can sometimes be observed to stop, say they have gone wrong, and then continue further in the wrong direction. Such errors are often accompanied by self-depreciatory remarks, which the tester has ample time to record. Many subjects seem to gain added enjoyment from the test after they have made an error, though this is less frequent in the case of the depressives.

There is a slight, but not significant, tendency for anxieties and depressives to enter a blocked passage more frequently in the first third of any maze than the other groups. They are also rather more prone to lift their pencils or to cross lines in the first two very simple mazes. The finding of Porteus that these errors are due to impulsiveness is difficult to reconcile with the present results (Porteus (6), 1942, p. 34).

Obsessionals. Rather slow tempo throughout, relative freedom from very gross errors of direction, few lifted pencils, firm, straight lines down the centre of the channel and careful right-angled cornering with consequently very few crossed lines. Chapuis (2) appears to be writing of normal people of a mildly obsessional character when he comments that 'Au T.L. le tracé anguleux est déjà en soi, dans la majorité des cas, un signe de propreté et exactitude, comme il révèle aussi fréquemment une nature simple et droite—sans énormément de souplesse' (p. 81). The low W.D. and W.W.D. scores may, in the present instance, be due in part to the higher intellectual capacity of this group.

Clearly the main differentiation on this test is between normals, psychopaths and hysterics on the one hand and anxiety states, reactive depressions and obsessionals on the other. The test does, however, show considerable differences in the pattern of performance among the clinical groups making up the dysthymic group. Differentiation between normals, psychopaths and hysterics is less satisfactory. It would appear, therefore, that poor Maze performance, under the present method of administration, is most closely linked with the presence of anxiety and depressive features.

Porteus (6) (1942) claims very considerable differences between normal children and delinquents and between normal adults and criminals. He appears, however, to have been content with only approximate matching for intelligence and to have accepted the Stanford-Binet test for this purpose. One would certainly expect that the Porteus Mazes correlated more highly with Progressive Matrices than with the Stanford-Binet. Close examination of his table (p. 19) suggests that the W.D. score of the bus drivers was very much lower than that of the criminals, though unfortunately he does not present all the information on this point. If this were so, and the usual correlation between Binet scores and W.D. scores is 0.6-0.7, it would be difficult to reconcile this result with even the approximate matching for intelligence which he claims.

VI. SUMMARY AND OUTLOOK

1. Certain modifications have been made in the administration of the Porteus Mazes with a view to decreasing the 'g' loading and emphasizing the temperamental factor involved in its performance.

2. An attempt has been made to assess the grossness of errors of direction as well as their number. The various categories have not been summed, as it is considered that the result would be psychologically meaningless.

3. Differences in performance on the Maze test due to sex, age, and psychoneurosis have been studied, and a comparison has been made between the results obtained from subjects suffering from the different forms of psychoneurotic illness.
4. The test was found to differentiate more clearly between the psychoneurotic groups than between psychoneurotics and normals.
5. The characteristic performances of psychopaths, hysterics, anxiety states, reactive depressions and obsessionals have been described.
6. Investigations are proceeding which, it is hoped, will go some way towards explaining the marked differences in tempo observed among psychoneurotic subjects.

REFERENCES

- 1) BABCOCK, H. (1941). *Time and the Mind*. Cambridge, Mass.
- 2) CHAPUIS, F. (1949). *Le Test du Labyrinthe*. Berne: Hans Huber.
- 3) DOLLARD, J. *et al.* (1944). *Frustration and Aggression*. London: Kegan Paul.
- 4) EYSENCK, H. J. (1947). *Dimensions of Personality*. London: Kegan Paul.
- 5) MEAD, M. (1949). *Male and Female*. London: Gollancz.
- 6) PORTEUS, S. D. (1942). *Qualitative Performance in the Maze Test*.
- 7) PORTEUS, S. D. (1946). Porteus Maze Tests in *Encyclopedia of Psychology* (ed. P. L. Harriman). New York: Philosophical Library.
- 8) PORTEUS, S. D. (1950). *The Porteus Maze Test and Intelligence*. Palo Alto, California.

(Manuscript received 16 October 1950)

A STUDY OF TEMPERAMENT IN ADOLESCENCE¹BY ALASTAIR HERON, *University of Manchester*

I. *Introduction* (p. 218). II. *The original observation and a resulting hypothesis* (pp. 218-219). III. *The fresh sample* (p. 219). IV. *Preparatory work* (p. 219). V. *Presentation of the test material* (pp. 219-223). VI. *Identification of two groups of boys and comparison thereof on all available criteria* (pp. 223-227). VII. *Interpretation of findings* (pp. 227-231). VIII. *Summary* (p. 232). *References* (pp. 232-233). *Appendices* (pp. 233-237).

I. INTRODUCTION

During the process of analysing the responses to a questionnaire on the general topic of education for parenthood, obtained from a sample of 699 adolescents aged 12-18, a small point was noticed which, though marginal to the main objective, seemed of potential importance. As a result, when the inquiry, first carried out in the schools of a predominantly rural county in south-west England, was repeated as planned in an industrial city of the north-west, special provision was made for the further investigation of this finding. The purpose of the present paper is to report the results of this investigation, and to offer some observations thereon.

II. THE ORIGINAL OBSERVATION AND A RESULTING HYPOTHESIS

The small point to which reference has been made arose in connexion with the distribution of responses by the boys' sample ($N=341$) to two of the statements in the questionnaire:

S.2: 'Boys are not interested in the care of children.'

S.6: 'Looking after young children in the home is a woman's job, and men cannot be expected to take much share in it.'

These responses were combined so that it was possible to see how many boys had *agreed* with *both* statements, and how many had *disagreed* with *both* statements. Table 1

Table 1

Age-group ...	12-13	14	15	16	17-18	Total
Agreed with both statements	15	27	20	13	10	85
Disagreed with both statements	17	20	17	19	10	83

shows the interesting pattern which emerged. These two contrasting groups, each representing about 25% of the boys' sample, were then examined to ascertain whether or not they differed on other available criteria. It was found that considerable differences between the two groups existed in connexion with the incidence of reported activities on behalf of younger children, and also of expressed dislike for such activities. The figures are given in Table 2. It would appear, therefore, that a relationship existed between expressed lack of interest in the care of younger children, reported lack of experience with them, and expressed dislike of such experience as they have had. No other information was available.

¹ The material presented in this paper formed part of a thesis, 'Cultural and psychological problems in education for parenthood', submitted as part of the requirements for the degree of Master of Science in the University of Manchester, June 1949. In highly condensed form, the main findings were presented under the thesis title at the XIIth International Congress of Psychology, held at Edinburgh in July 1948. The writer wishes to record his gratitude to Dr F. E. MacNeill for her advice and criticism throughout this research, and to Mr L. T. Wilkins for his help on various statistical points.

Table 2

	<i>N</i>	Activities reported	Disliked as % of total
Agreed with both statements	85	200	19.5
Disagreed with both statements	83	308	7.7

but even these scanty findings suggested a possible hypothesis, which may be stated formally as follows:

Adolescent attitudes to younger children are a manifestation of personality differences which are by then relatively independent of cultural norms.

III. THE FRESH SAMPLE

On this occasion it was possible through the co-operation of the Director of Education¹ to carry out a most satisfactory sampling procedure. Access was given by him to the duplicate registers of all the schools in the Local Education Authority's jurisdiction which contained children of 12 years of age and older. The total population of boys and of girls in each of the six age-groups 12 to 17 + inclusive was first ascertained. It was then possible to allocate to each school a quota of the age-group samples proportionate to the contribution which the school made to the total age-group population of the area. The sample for each age-group was twenty. Within each school the number of individuals required for the school's quota was distributed at random, first as to the classes containing boys or girls of the required age, and secondly within the classes thus randomly selected. All these selections were carried out using a table of random numbers. This entire procedure was carried out twice, thus providing two independently drawn samples each consisting of 120 boys and 120 girls (Table 3). These two independent samples were compared for homogeneity on the objective criterion provided by the distribution of family size (Table 4). It will be seen that homogeneity may be inferred with an extremely high degree of probability, and we may therefore conclude the sampling to have been adequate. This conclusion is reinforced indirectly by such other criteria as the distribution of fathers' occupations and by a comparison for homogeneity of the two samples on certain of the psychological tests employed.

IV. PREPARATORY WORK

The writer was introduced to each head teacher by a letter from the Director, or, in some cases, personally. At a preliminary interview, and subsequently, it was stressed that no word as to the nature of the inquiry should reach either teachers or pupils. The object of this precaution was to ensure (a) that the presentation of the material was standardized, and (b) that no thought or discussion on the subject took place before the writer's visit. This objective, was, so far as he is aware, satisfactorily attained.

V. PRESENTATION OF THE TEST MATERIAL

At the prearranged time the writer presented himself at the school, and handed the head teacher a list of the children selected by the random procedure described above; if any child was absent, a substitute was available from a pool of the same age also randomly selected. This scheme worked very smoothly and enabled the work to be done with a minimum of disturbance to the running of the schools. Generally speaking, the group of children required in each school was assembled in an empty classroom, but where this was not

¹ For the County Borough of Oldham, Maurice Harrison, M.Ed., to whom the writer is most grateful.

Table 3. *Showing the composition of the whole sample after combining the two independently drawn half-samples ($N = 2 \times 240 = 480$)*

Age-group	...	12		13		14		15		16		17 +		Totals		
		Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Total
Type of school		Pop.	S.	Pop.	S.	Pop.	S.	Pop.	S.	Pop.	S.	Pop.	S.	Pop.	S.	
Secondary Modern		359	24	387	26	388	24	342	22	357	24	—	—	—	—	72
Church of England		75	4	63	4	63	4	52	4	73	6	—	—	—	—	10
Roman Catholic		82	6	66	4	75	4	78	6	68	4	—	—	—	—	14
High Schools		91	6	102	6	131	8	88	6	169	10	73	6	89	40	144
Totals		607	40	618	40	640	40	559	40	631	40	561	40	89	40	240

Notes: (1) 'Pop.' figures indicate the total population of children of that age and sex in all the schools of the specified type.

(2) 'S.' figures indicate the quota of the age-group sample allocated to that type of school.

(3) 'High Schools' include Junior Art and Technical Schools.

(4) At the 17 + age-level, the schools of the L.E.A. were unable to complete the age-group sample, and it was therefore necessary to make up the required number by drawing 11 boys and 20 girls at random from the nearest available school.

(5) The sample was drawn from 22 schools.

Table 4

Sample	Number of children in family										
	1	2	3	4	5	6	7	8	9	10	11
I	60	70	54	23	14	7	6	5	—	—	1
II	53	84	49	21	14	10	7	2	—	—	—
Totals	113	154	103	44	28	17	13	7	—	—	1

$\chi^2 = 1.629$ (D.F. 8); $P = 0.99$.

possible, a corner of the school hall or even the head-teacher's office was used. He was then introduced by the head teacher to the group (or the first group, in the case of a school with a large quota), using close approximations to the sentence, 'This is Mr Heron, who has some work he would like you to do for him'. Teacher and head teacher then left the room. The investigator then explained to the class that he was neither teacher, inspector, nor examiner; that he did not even require their names on the papers; and that this work was being done in many schools. In every case, this introduction effected good rapport consequent upon relief from anxiety about his identity, and distribution of the material followed.

The material consisted of the following sections, each of which the whole group completed before passing on to the next section. Every item of the background information section was explained separately by the writer, and detailed overall explanations were provided for every other section, with opportunity for questions.

(a) *45-item vocabulary test.*

From form 'L' of the Terman-Merrill Revision of the Stanford-Binet Intelligence Test. The children were told to 'write down simply what the word means. For example, an "envelope" is "what you put a letter in". IS THAT CLEAR? Now when I tell you to start, I want you to do *as many* as you can *as quickly* as you can; don't worry about the spelling, it doesn't matter for this. Don't waste too long on any one of them, come back if you have time. You have *ten minutes*: START NOW!'

This ten-minute period was rigidly adhered to. Children who were found to be unable to write, or in some cases to read, at the age of 12, were dealt with individually and their responses to the questionnaire as a whole recorded by the investigator. This was made possible by the small size of the groups examined and occasioned no difficulty whatever. An examination of the means and standard deviations from each age-group confirmed an earlier impression that this use of the vocabulary would serve as a *rough* test of capacity to deal with verbal material, but it was used in this stage of the work as a means of comparing groups rather than for selecting individuals.

(b) *Background information inquiry* (see Appendix I).

It was repeated here that, as their names were not required, they could freely provide the information required. They were warned NOT to state *where* their father or mother worked.

(c) *Activities for younger children* (see Appendix I).

(d) *Twelve-statement questionnaire on education for parenthood* (see Appendix II).

Two forms of this were prepared, one being as nearly as possible the inverse of the other; in other words, statements in the positive on form A appeared in the negative on form B, and vice versa. From previous experience, the statements were also balanced internally so that one could anticipate a majority of 'agree' responses to half the statements. The theoretical basis for this technique may be explained briefly as follows: In distributing the complete sets of papers, it was arranged that the two forms of the twelve statements were issued alternately, with the result that *for each form*, in each age-group, for each sex, there were twenty *sets* of responses. The total numbers of agree/disagree/undecided responses for all the twelve statements in each of these groups of twenty could then be compared. In making the comparison it is of course necessary to remember that agreement in form A in respect of any given statement is equivalent to disagreement to form B of the same statement and vice versa. As in both forms there are six affirmative and six negative statements, and as the expectation of majority response for agreement and disagreement has been approximately equalized, an hypothesis may be set up: *If for any given age-group the subject-matter of the twelve statements is on the whole meaningful, responsibly dealt with, and relatively free from suggestion, the distribution of agree/disagree/undecided totals for each form will not differ significantly.* A subsidiary hypothesis is, of course, that this criterion will not be met by the younger boys and girls. Combining the two hypotheses, it was suggested that by this method we should be able to establish objectively the minimum age at which, on the average, this sort of material can be dealt with meaningfully by boys and girls respectively.

An examination of the figures in Table 5 will show that these hypotheses were confirmed. It is at once evident that the girls are *at least* one year later in meeting this criterion than are the boys, who do so by the beginning of the 15th year. This may come as somewhat of a surprise to those who assume that the

nature of the subject-matter would ensure just the opposite result. What is overlooked in this assumption, however, is the tendency towards greater suggestibility and a relatively more emotional reaction in adolescent girls than in boys(26).

Table 5. *Comparison of total agree/disagree/undecided responses to the twelve statements in their alternative forms*

For each form, in each age-group, for each sex, $N=20$; therefore the total responses A. + D. + U. in each cell thus formed = $12 \times 20 = 240$.

Age-group	...	12			13			14			15			16			17 +		
		A.	D.	U.	A.	D.	U.	A.	D.	U.	A.	D.	U.	A.	D.	U.	A.	D.	U.
Boys																			
A		120	70	50	102	71	67	111	78	51	106	97	37	110	93	37	112	95	33
B		73	110	57	64	121	55	100	100	40	101	105	34	93	109	38	104	116	20
χ^2		20.792			22.899			4.622			0.564			2.704			5.575		
(D.F. 2) P=		0.001			0.001			N.S.			N.S.			N.S.			N.S.		
Girls																			
A		127	67	46	128	54	58	126	76	38	133	82	25	98	102	40	99	108	33
B		56	102	82	75	110	55	87	92	61	104	101	35	101	101	38	112	104	24
χ^2		44.919			33.039			14.008			7.187			N.S.			N.S.		
(D.F. 2) P=		0.001			0.001			0.001			0.05			.091			2.297		

(e) 125-word *Pressey-Collins interests list* (see Appendix III).

The first requirement for the investigation of the special topic now under review was to provide a means whereby the whole sample of boys could be divided roughly but objectively into those whose responses to children could be described as positive and as negative. It occurred to the writer that what was needed was a list of words or sentences which contained references to 'babies' or 'children', and preferably to both. A further requirement was that these references should be very adequately 'buffered'. The ideal instrument for these purposes seemed to be the 'Interests' section of the Pressey 'X-O Test' as modified by Collins(7), which had the further advantage of being serviceable in its own right as a measure of relative 'affectivity' and also of providing a means whereby constellations of interests connected with the two groups of boys might be profitably identified.

(f) *Domestic situations and tasks* (see Appendix IV).

It also seemed necessary to provide some means whereby it could be established whether the norms for domestic situations and the behaviour appropriate to them differed significantly as between the two groups. This was met by adaptations of portions of Burdick's Group Test of Home Environment(5).

(g) *Pity test* (see Appendix V).

In considering the possible nature of the variables which might be involved in this difference it was perhaps not surprising that attention should be paid to that of 'relative masculinity'. The writer therefore consulted Terman & Miles's(25) monumental work on this subject, with a view to making use of one or other of their more successful subtests. The most appropriate according to their own assessment appeared to be the 'Pity test', which forms part of their section on 'Emotional responses'. Of this the authors say (p. 36): 'The emotional-attitude test proved successful beyond all expectations. . . . The reliabilities were surprisingly high even when computed separately for the four parts of the test.'

The reliability for the Pity section is given as 0.86 for the twenty-nine items which they used. From these twenty-nine items the writer selected ten from among those which they report as showing the greatest discriminatory value between the sexes. Of these ten, three (nos. 1, 2 and 6) have been modified slightly.

Scoring.

The tests listed above have been scored as follows:

(a) *Pressey*. As a measure of 'affectivity' the individual's score is represented by the total number of words underlined. It was found that no practical use could be made of the most favoured item in each line.

(b) *The domestic situations*. An analysis was made of the verbatim responses of a stratified subsample consisting of ten boys and ten girls from each of the six age-groups, which is equivalent to 25% of the

otal sample. It was found that the responses to each of the eight situations could be categorized after dividing them into three topics:

Discipline (nos. 1, 2, 3, 4, 5 and 8);

Maternal co-operation and interest (no. 6);

Home hospitality for close relatives (no. 7).

The categories, six in number, are exemplified as set out in Table 6. The responses of each individual to each situation were then marked with the number of the category into which they fell, and from the individual data distribution tables were compiled for the two groups of boys and for the girls. This avoids the application of a subjective rank-ordering of the categories on the part of the writer, such as would be necessary if an attempt was made to provide each individual with a score.

Table 6. *Classification of responses to the 'Domestic Situations Test'*

I	II	III	IV	V	VI
		<i>Discipline: items 1, 2, 3, 4, 5 and 8</i>			
Reasoning	Non-physical punishment	Bed	No food	Smacking	Weakness
Explaining			Shouting	Hitting	'Remarks'
Firmness			Threatening	Clouting	No ideas
Help	'Very angry'		Frightening	Hiding	No response (from child)
Interest			Bribing	Thrashing	
Encouragement	Scolding		'Bad'	Whipping	
Mediation			Swearing	Walloping	'Tell father'
Reconciliation	Telling-off		Sarcasm	Belting	
Understanding				Spanking	
	'Lecturing'			Beating	
				Chastising	
				Strapping	
		<i>Maternal co-operation: item 6</i>			
Encouragement	Model	Read	Play	Radio	As above
Interest	Paint		Wash		
Help	Draw	'Tell each other stories'	Eat		
Join in	Make				
	Do things				
		<i>Home hospitality: item 7</i>			
Musical evening at home	Have a party at home	Play cards at home	'Pub'	Cinema	As above
			Theatre		
			Concert		
			Show		
			Walk		

(c) *Twenty domestic tasks.* Individual responses for this test were totalled under two headings: tasks for which the names of *males* had been recorded, e.g. 'father', 'brother', 'self' (if a boy), and those regarded as appropriate for 'the children'. In addition, complete distributions were compiled under the five headings 'mother', 'father', 'boys', 'girls', 'children' in order to permit direct comparison between boys and girls on what they regarded respectively as the norm in such matters. These distributions show that boys and girls both reach their maximum readiness to respond to this sort of material in their sixteenth year, and confirms that the differences between them are minimal during the same year.

(d) *Pity test.* This was scored as indicated by Terman & Miles (25).

VI. IDENTIFICATION OF TWO GROUPS OF BOYS AND COMPARISON THEREOF

ON ALL AVAILABLE CRITERIA

As indicated above, the presence of the words 'babies' and 'children' in the Pressey test was used as a means of dividing the boys' sample ($N = 240$) into those who underlined either or both of these words, and those who underlined neither. The result was as shown in Table 7,

Table 7

	Boys							Girls						
Age-group...	12	13	14	15	16	17 +	Total	12	13	14	15	16	17 +	Total
Sample I	11	8	5	10	11	7	52	16	18	15	20	16	12	97
Sample II	7	7	6	9	9	10	48	17	17	16	18	15	20	103
Did	18	15	11	19	20	17	100	33	35	31	38	31	32	200
Did not	22	25	29	21	20	23	140	7	5	9	2	9	8	40
Total	40	40	40	40	40	40	240	40	40	40	40	40	40	240

which also gives the comparable figures for girls. The χ^2 test indicates homogeneity of samples I and II, and also that age-group distributions do not differ significantly from chance expectation.

It will be seen that

- (1) Just twice as many girls as boys expressed interest in or liking for babies or children;
- (2) That the total figures for the two independent samples referred to earlier when describing the sampling procedure are almost identical both for boys and for girls.

From now on all references to the *positive* group of boys should be identified as to those *who expressed interest in babies or children*. The *negative* group therefore refers henceforth to those who did *not* do so. These two groups, and the complete sample of girls, were compared on all the criteria designed for the purpose, with the addition of one or two others which, during the analysis of the data, seemed likely to be useful. This comparison is set out in full in Table 8, to which attention is now directed. It will be seen that the figures have been grouped so as to show together, on the one hand, those criteria upon which there was *no* significant difference between the two groups of boys, and, on the other hand, those upon which there *was* a significant difference. From this it may be concluded that the two groups of boys did not differ significantly on criteria which are concerned with local cultural norms, the presence or absence of younger brothers or sisters, or of their mothers, their intelligence (in so far as this can be said to be measured roughly by the vocabulary group test), nor upon their responses to those items of the statement questionnaire which are not specifically concerned with their interest in young children. The two groups do differ significantly on criteria connected with what may perhaps be most conveniently described as the orectic aspects of personality. The original hypothesis—that adolescent attitudes to younger children are a manifestation of personality differences which are by then relatively independent of cultural norms—may be regarded as sustained.

In the light of the foregoing information about the two groups it seemed desirable to carry out a detailed analysis of the responses to the Pressey test, with a view to establishing whether or not the two groups of boys could be shown to differ in certain of their interests. This decision is in accord with Murray (21).

In distinguishing these average people, focalities (enduring cathexes) are as important as generalities (needs). This fact immediately suggests a type of questionnaire... (which) could be modelled on some of the already devised attitude or annoyance questionnaires. It would consist, let us say, of a long list of commonly liked and disliked objects or situations which the subject would be asked to score on an affective basis.

The method adopted was as follows:

(a) The age-group containing 15-year-old boys of both groups separately, and of girls was selected as being the one in which differences might be expected to be minimal (and therefore more significant);

(b) For each of the three groups, 'positive', 'negative', 'girls', the total number of times each of the 125 words in the test had been underlined was ascertained;

(c) For each group, the 125 words were then placed in order of popularity and rank-order correlation coefficients calculated; these were as follows:

'Negative' boys/'Positive' boys:	+0.824 (P.E. $\pm$ 0.03).
'Positive' boys/girls:	+0.687 (P.E. $\pm$ 0.05).
'Negative' boys/girls:	+0.541 (P.E. $\pm$ 0.06).

This result may be regarded as supporting the general hypothesis that the 'positive' group of boys have more things in common with the girls than has the 'negative' group. Obviously all 15-year-olds have *some* things in common, regardless of sex or temperament.

Table 8

Criterion	a	b	c	d	e		f	g	h	i	j	k	l	m
	Younger siblings %	Mothers working %	Male tasks Mean	Vocabulary Mean	S.: 2, 3, 5, 7, 8, Domestic 9, 10, 11 and 12	Distr.	Distr.	Pressey Mean	Pity Mean	Activity for younger children Rate per 100	Dislike of activity %	Agreement with S.4 %	No Church affiliation %	Interest in prayer %
Basis of comparison														
Girls (whole sample)	50.0	42.5	7.30	19.36	—	—	—	51.25	6.171	566	4.6	34.6	16.2	44.4
'Positive' boys	50.0	41.0	9.50	19.20	—	—	—	56.40	7.550	405	7.9	18.0	11.0	41.0
'Negative' boys	50.0	41.0	9.03	19.01	—	—	—	42.50	9.036	318	16.2	41.0	23.0	23.0
χ^2 or 't', positive/negative	—	—	—	—	—	—	—	6.944	2.531	—	12.772	14.410	4.799	7.686
P =	—	—	—	—	—	—	—	0.001	0.02	—	0.001	0.01	0.05	0.01
	No significant difference							Differences significant						
'Positive' boys in extreme position	—	—	—	—	—	—	—	×	—	—	—	×	×	—
'Positive' boys intermediate between 'negative' boys and girls	—	—	—	—	—	—	—	—	×	×	×	—	—	×

Notes: (1) χ^2 was used to test all figures shown as '%', and 'Distr.', except item (i), which does not lend itself to any test of significance. In 2 x 2 distributions, Yates's 'Correction for continuity' was applied.

(2) The 't' test was used for all figures shown as 'Mean'.

(d) A list was made of the words which in each *pair* of groups had occasioned the greatest rank-order difference, with the provision that no word was used for discriminatory purposes which did not show a difference of 25 or more rank-order places in its popularity with the two groups being compared. (The adequacy of this arbitrary lower limit of reliable discrimination was checked for the two boy groups by ascertaining the fiduciary limits of the mean of the rank-order differences at the 1% level. Those limits were 12 and 17; it will be seen, therefore, that the arbitrary limit of 25 referred to above ensures that we only employ those words which are well above average in their contribution to the total rank order difference. Because of the small size of the two groups, however, this does not constitute a guarantee that every difference will differ significantly from chance probability. So long as we are not desirous of making interpretations based upon single words, this is not a matter of great importance, bearing in mind the legitimacy of summated probabilities.¹

Table 9

'Positive' group	A. <i>Objective criteria</i>	'Negative' group
High	Affectivity (Pressey score)	Low
High	Emotional attitude (Pity score)	Low
High	Prayer interest (Pressey item)	Low
Low	Non-affiliation church (background data)	High
Low	(S.A. agreement) 'Boys are not interested in young children'	High
Low	Dislike of activities for young children	High
High	Activities for young children	Low

B. From rank-order differences in Pressey test at age 15. Relatively more interest in:

Children	Arguing
Babies	Napoleon
Grocers	Flirting
Teachers	Smoking
Banquets	Racing
Mushrooms	Chauffeurs
Hymns	Revivals
Good boys	Hotels
Church	Athletic girls
Circuses	Mowgli
	Galahad

C. Words obtaining majority in one group, minority in the other

Teachers	Arguing
Hymns	
Good boys	
Church	

D. Subjectively-produced descriptive material from 14 adults based solely on inspection of the bi-polar constellations of interests shown above

Cautious	Adventurous
Femaleness/objects	Maleness/activity
Victorians	Moderns
Home-makers	History-makers
Domestic	Dramatic
Family	Good time
Down-to-earth	Materialistic-idealistic
Security-seeking	Insecurity-wishing
Peace-seeking	Conflict-seeking
Prosaic	Romantic
Social inside	Social outside
Unexciting/inactive	Exciting/active
Feminine	Masculine
Cissy	Tough

¹ The writer subsequently discovered that a similar method had been used for other purposes by Chambers, *et al.* of Pressey's early collaborators. Cf. *Pediat. Seminary*, 32, 1925, 637-47; and *J. Abnorm. Soc. Psychol.* 20, 1925, 303-11.

The three lists of words thus obtained are therefore those which showed the greatest discriminatory value. Each list of words can be divided into two parts according to the *direction* of the rank-order difference for each word in the list. This provides, for each inter-group comparison, two lists of words; that shown for each group is the list which the group ranked significantly higher in each case than did the other group with which it is being compared. This procedure provides us therefore with a pair of polar constellations for each inter-group comparison.

Let us be quite clear, however, about the precise nature of these lists of words. They are those which, on the basis of their *relative popularity within each group*, display the greatest differences between the groups. This is *not* the same as saying that each list consists of those things which each group respectively holds most in favour, for when comparing two groups of boys of the same age, these *most popular* things obviously tend to be the same things. This is clearly evidenced by the rank-order correlation coefficient of 0.824 cited above. This being so, we must recognize that each list of words is an abstraction, a rather shadowy 'gestalt'; while it is certainly not without value to our purpose, we cannot employ it except in conjunction with other data. Used in this way, however, it can contribute in its turn to the total pattern of information about the groups being compared.

It was seen at once that the lists obtained from comparing both groups of boys with the girls provided characteristically boy/girl constellations. This justified our contemplation of the lists resulting from comparing the two groups of boys *with each other*, with reasonable confidence in the significance of the differences between them (section B, Table 9).

If to these two lists of words the further criterion is applied, that only those words shall be considered which obtained a *majority* vote from one group and a *minority* vote from the other, then two briefer lists appear, as set out in section C of Table 9.

It is common practice in psychology, whenever an investigator is forced by the nature of his material to make subjective evaluation, for him to seek similar evaluations from others as a form of necessary control. By adhering to this principle on this occasion, it was possible to obtain a sufficiently broad base of description for the two groups to assist the writer in his attempt to compare them with existing typologies. This descriptive material is set out in full as section D of Table 9.

VII. INTERPRETATION OF FINDINGS

Before proceeding, one obvious methodological point requires attention. It had seemed necessary to plan this investigation so as to provide for the individual interviewing of a 10% subsample of the boys at the conclusion of the large-scale group inquiry. This, however, could only have been made possible by violating the anonymity of the boys responding to the questionnaires. For a number of reasons it was decided that the guarantee of anonymity would have to stand. In view of the results already presented, it is indeed unfortunate that this decision had to be taken, but if the work was to proceed at all, there was no alternative.

We are therefore obliged to attempt an interpretation of the differences which have been shown to exist between these two categories of boys, as set out in Table 9. The obvious temptation is to accord each group of boys a descriptive type name, and it was only with difficulty that this temptation has been avoided during the latter stages of this investigation. Dichotomous typologies abound in the psychological literature—a good example of

an up-to-date list is to be found in *Dimensions of Personality* (9). One may quote with approval Eysenck's remark (p. 14): 'In spite of obvious similarities, there are also profound differences between any two of them, and much confusion of thought has resulted from a glossing over of these differences.' In these circumstances there is an understandable reluctance to suggest yet another pair. It seems more profitable to examine the characteristics of our two groups with a view to establishing the extent to which they may be adequately explained by one or more of the existing typologies. A careful examination was therefore made of the modes of description employed by the authors responsible for the literature concerning the more likely bi-polar continua. The reader is now invited to consider the italicized portions of the following selection of extracts, and to compare them with the present findings.

(1) James (13) includes *religious* and *idealistic* in his 'tender-minded'; and Jung (14), commenting on James's famous dichotomy, says: '*tender-minded* has certain traits in common with the *realist*... and realism corresponds with the principle of *introversion*...'

(2) Ostwald (23) is quoted by Jung (pp. 401 ff.) on the former's *classic* and *romantic* types: 'While the former is characterized by the well-rounded perfection of each individual achievement, and at the same time by a *rather withdrawn nature* whose personal effect upon his environment is but slight, the *romanticist* stands out by reason of the very opposite characters... *Pioneers* who possess great reactive rapidity are the *romantics*, while those with slower mental reactions are the *classics*... *a never-failing characteristic of the classic type is the absolute need to stand without error or blemish in the public eye.*'

Jung himself continues (pp. 403 ff.): 'As one reads the Ostwald biographies, one sees at once that the *romantic* type corresponds with the *extravert*, the *classic* with the *introvert*... The outward reaction is characteristic of the extravert, just as the inner reaction distinguishes the introvert.' '(The introvert) suppresses his reactions, which, however, can be just as quick as those of the extravert—hence the introvert may easily give the impression of slowness... the *romantic* is certainly able to *display his pre-maturity*, but the classic although perhaps equally mature, may conceal his products within himself, not designedly of course, but from an inability for immediate expression. As a result of deficient differentiation of feeling, the introvert exhibits a certain awkwardness, a real infantilism in the personal relation, i.e. in that element which the Englishman calls "personality". His expression, his "personality", *appear commonplace to the superficial judgment*, while not infrequently the *romantic immediately appears "interesting"*, and understands the art of encouraging this impression by every sort of means, *whether permissible or not.*'

(3) Oliver (22), working only with extreme extraverts and introverts, mentioned 'a definite tendency for *introverts* to favour *stricter moral standards*'.

(4) Murray (21) elaborates his general reference to the psycho-analytic concept of the 'super-ego' by describing what he calls the '*positive super-ego*' as follows: 'It presents to the individual certain *ideals of social or saintly conduct*, the conception of a life consecrated to mankind or God.' This usually involves (various) forms of the 'need Deference (obedience) and the 'need Nurturance' (charity). Of the latter he remarks, 'To give sympathy and gratify *the needs of a helpless object*', and quotes as typical actions, '*to enjoy the company of children and animals*'; trait-names and attitudes appropriate to the 'need are '*gentle, maternal... protective... paternal*'.

In this connexion Freud (11) describes the 'establishment of the super-ego... as *successful instance of identification with the parental function*'.

(5) Terman & Miles⁽²⁵⁾, in the course of their discussion of *masculinity-femininity*, say of the typical woman in occidental cultures that she 'is believed to differ from the typical man in the *greater richness and variety of her emotional life* and in the extent to which her very-day behaviour is emotionally determined. In particular, *she is believed to experience a greater degree than the average man the tender emotions, including sympathy, pity and parental love*; to be more given to cherishing and protective behaviour of all kinds. Compared with man she is more timid and more readily overcome by fear. *She is more religious* and at the same time more prone to jealousy, suspicion, and injured feelings. Sexually she is by nature less promiscuous than man, is coy rather than aggressive, and her sexual feelings are less specifically localized in her body. *Submissiveness*, docility, inferior steadfastness of purpose, and a general lack of aggressiveness reflect her weaker conative tendencies. Her moral life is shaped less by principles than by personal relationships, but thanks to *her lack of adventurousness* she is much less subject than man to most types of criminal behaviour.'

When considering the masculinity-femininity concept, however, it is necessary in the light of anthropological data^(15, 16, 19, 20) to examine the possibility that such a concept is culturally relative. This can be illustrated in the present instance by asking ourselves quite simply, 'What do we mean when we describe the boys who like children as "more feminine"?' The answer is, of course, that in certain of their characteristics they approximate to what is understood in *our culture* as a 'feminine' reaction-pattern. To quote Freud⁽¹¹⁾, p. 147): 'We speak of a human being, *whether male or female*, behaving in a masculine or feminine way.... The difference is not a psychological one; when you say "masculine" you mean as a rule "active", and when you say "feminine" you mean "passive".'

Similarly, Maslow, as quoted by Boring, Langfeld & Weld⁽⁴⁾, p. 513), said of monkeys that 'regardless of gender, the *dominant* animal played the *masculine* role in sex activity, initiated *fighting* and play and was more active. The *submissive* animal played the *feminine* role....'

(6) The Guilfords'⁽¹²⁾ 'M' factor (masculinity-femininity) is referred to by G. W. Allport⁽¹⁾ as 'a vague factor confused with the altogether irrelevant consideration of sex. As much as anything it is an empirical counterpart of the more orderly variable described previously as "*ascendancy-submission*".'

From such a comparison two things stand out with considerable clarity:

(a) our two groups of boys present characteristics suggestive of each of the bi-polar continua which we have quoted; and

(b) because they do so, they suggest the possibility of integrating these various theoretical and empirical concepts in a psychologically meaningful way.

There seems little doubt that from a variety of approaches, and subject to a general recognition that the concepts are over-inclusive, Eysenck, Murray, Allport and the Guilfords have tacitly accepted the 'extraversion-introversion' formulation of Jung. The analyses of Murray and the Guilfords do much to explain why so few people are found who present *all* the general characteristics of 'extraversion' or 'introversion', and Allport, Murray and Burt⁽⁶⁾ are emphatic in regarding both concepts as independent and positive, thus providing for their simultaneous operation within a single personality. This approach extends and clarifies Jung's original, and rather confusing, suggestion that introversion was operative unconsciously in the extravert, and vice versa.

If we apply this synthesis to our earlier impressions and data concerning the two groups of boys, it can readily be seen why the respective 'gestalten' gave rise to various characterizations shown to be subsumed under the extraversion-introversion concept. But we have also seen that, so far as these two groups were concerned, this concept was far too general, in view of the more specific tendency for the description to be in terms of 'tough and tender', 'masculine and feminine'. Our attention must therefore be directed to those specific aspects or manifestations of 'extraversion' and 'introversion' which seem likely to lead us towards a psychologically meaningful explanation of the data.

It seems reasonable to suggest that the traits 'ascendancy' and 'submission' are probably adequate to describe what our 'Western' culture would regard as the most outstanding and appropriate features of 'masculine' and 'feminine' behaviour respectively. For our purpose, therefore, they are the focus of attention, as in view of the evidence from anthropology we must seek descriptions and explanations which are *psychological* in reference; that is, free from the relativity inherent in culturally approved words and phrases. The group of boys interested in children, then, might be described provisionally as *characterized by an excess of 'submission' over 'ascendancy'*; but it must be re-emphasized that this may reflect a positive style of life, a pattern of adjustment—it is not necessarily 'weak' *psychologically*. Conversely, a style of life characterized by an excess of 'ascendancy' over 'submission' may not indicate psychological 'strength'. Bearing this constantly in mind, we may now turn to the study of the psycho-genesis of the 'submissive' trait. According to Murray, it may be analysed into four 'needs', which he styles: 'inferiority-avoidance', 'blame-avoidance', 'seclusion' and 'abasement'. The second and fourth of these are also included by him in his analysis of the negative functions of the Freudian 'super-ego'. He also refers to the positive functions of the 'super-ego', as already quoted, and to their connexion with his 'need Nurturance'. The trait-names which he suggests for this 'need' include 'gentle', 'maternal', 'protective' and 'paternal', and he includes among actions typical of its presence, 'to enjoy the company of children and animals'. In describing the positive functions of the 'super-ego', he says: 'It presents to the individual certain ideals of social or saintly conduct, the conception of a life consecrated to mankind or God.' If Murray's analysis is accepted and applied to our data and previous reasoning, the result is clearly that our group of boys who are positive towards children may be described as *'characterized by a well-developed "super-ego", both in its positive and negative functions'*.

Will such a description stand when the case of the *girls* is taken into consideration? The relevant questions are:

(a) If the presence in *boys* of an excess of 'submissiveness' (resulting in a reaction-pattern regarded as 'feminine' in our culture) is to be identified theoretically with a well-developed 'super-ego', are we justified in saying the same about the girls who exhibit the same trait?

(b) Are there any girls who do *not* exhibit an excess of 'submissiveness' as reflected in their reaction-pattern on our present criteria? If there are, is the proportion the same as that of the boys?

(c) If not, why not?

The answer to (a) depends largely on those to (b) and (c).

It will be remembered that the criterion devised for separating the two groups of boys was the expression of interest in, or liking for, 'babies' or 'children' or both, by means of

the Pressey interest test. When the figures for boys were reported (Table 7), those for the girls were also quoted for comparison: the totals may be set out afresh here:

	Boys	Girls	
Expressed interest	100	200	(Positive groups)
Did not do so	140	40	(Negative groups)
Totals	240	240	

Despite the smallness of the numbers involved in the 'negative' group of girls, let us examine the comparative responses on the three main criteria which proved significantly discriminating for the boys (see Table 10). These figures provide an effective and convincing answer

Table 10

	Girls		Boys	
	Positive	Negative	Positive	Negative
'Affectivity' (Pressey mean)	53.5	40.0	56.4	42.5
Non-affiliation to church (%)	13.0	32.5	11.0	23.0
Dislike of activities for young children	3.4	11.0	7.9	16.2

to question (b), but we are left the more difficult task set in question (c): Why are there three and a half times as many boys as girls whose expressed attitude to children is negative? To this one is inclined to answer in terms of cultural pressure, and without much more adequate information about the individual boys and girls, there is little point in speculating further. One suggestion, however, seems legitimate: were it not for the cultural pressure, the proportions of boys and of girls in each group would tend towards equality. The position can be restated thus. In spite of the continuous pressure of the culture upon the girls in respect of their 'feminine' role from early childhood onwards, there are *still some who each adolescence negative in their attitude towards children*, and this attitude is associated significantly with *the same criteria which characterized the differences between boys*. Positive evidence is lacking for the girls equivalent to that which directed attention to the 'super-go' in the case of the boys,¹ but it does not seem to be an unjustified inference to answer question (a) in the affirmative.

If this is agreed to, however tentatively, we may summarize the interpretation of the present data as follows: Both in boys and in girls an expressed interest in young children tends to be associated with evidence of an excess of the trait of 'submissiveness' over that of 'ascendancy', in so far as these traits are recognized as unitary variables contributing psychologically to 'introversion-extraversion' and culturally to 'femininity-masculinity', and that this adolescent situation is probably connected with the strong development of that aspect of personality styled by psycho-analytic writers the 'super-ego'.

It is indeed unfortunate that it was not possible to carry out the original plan to interview individuals in each group, even on a sampling basis; but definite and wholly objective findings remain, based upon an adequate and representative sample of adolescents. Interpretations of such findings may be expected to differ, but it is hoped that the one which has been offered may provide a basis for fresh hypotheses in connexion with developmental psychology, in so far as it may suggest an integration of related concepts in this field.

¹ i.e. the lists of significantly discriminating interests; the numbers in each age-group of 'negative' girls are too small to permit the use of this procedure.

VIII. SUMMARY

(a) It has been clearly established that within the age range 12 to 17+ boys may be divided significantly into two groups: those who do and those who do not express interest in young children.

(b) Those two groups do *not* differ significantly on criteria associated with

- (i) the presence or absence of younger siblings;
- (ii) intelligence, as estimated roughly by means of a group vocabulary test;
- (iii) local norms concerning family behaviour;
- (iv) reactions to the general proposal of education for parenthood.

(c) These groups *do* differ significantly on criteria connected with *activities for young children* and *with the orectic aspects of personality*.

(d) While in some respects those positive in their attitude to young children tend to occupy a position intermediate between boys negative in their attitude and the girls, in other respects the positive group is found in an 'extreme' position which is on the opposite side of the girls' sample from that occupied by the 'negative' boys.

(e) Evidence was also found that similar distinctions may be made among *girls* in the same age-range, although the proportion who do not express interest in young children is considerably smaller than is the case with the boys.

(f) An attempt has been made to relate the present findings to existing concepts in the field of personality, and as a consequence to provide a tentative basis for their integration

REFERENCES

- (1) ALLPORT, G. W. (1937). *Personality: a Psychological Interpretation*. New York: Holt.
- (2) BLACKBURN, J. (1945). *Psychology and the Social Pattern*. London: Kegan Paul.
- (3) BLACKBURN, J. (1945). *The Framework of Human Behaviour*. London: Kegan Paul.
- (4) BORING, LANGFELD & WELD (1935). *Psychology*. New York: Wiley.
- (5) BURDICK, E. M. (1928). A group test of home environment. *Arch. Psychol.*, N.Y., CI.
- (6) BURT, C. (1940). *The Factors of the Mind*. London: University of London Press.
- (7) COLLINS, M. (1928). British norms for the Pressey cross-out test. *Brit. J. Psychol.* XVIII, 121-33.
- (8) EDWARDS, A. L. (1946). *Statistical Analysis for Students in Psychology and Education*. New York: Rinehart.
- (9) EYSENCK, H. J. (1947). *Dimensions of Personality*. London: Kegan Paul.
- (10) FISHER, R. A. & YATES, F. (1948). *Statistical Tables*, 3rd ed. Edinburgh: Oliver and Boyd.
- (11) FREUD, S. (1946). *New Introductory Lectures on Psycho-Analysis*, 3rd ed., trans. by Sprott. London: Hogarth.
- (12) GUILFORD, J. P. & R. B. (1936). Personality factors S, E and M and their measurement. *J. Psychol.* II, 109-27.
- (13) JAMES, W. (1907). *Pragmatism*. London: Longmans, Green.
- (14) JUNG, C. G. (1924). *Psychological Types*. London: Kegan Paul.
- (15) KARDINER, A. (1945). *The Psychological Frontiers of Society*. New York: Columbia Univ. Press.
- (16) KLUCKHOHN, C. & MOWRE, O. H. (1944). Personality and culture: a conceptual scheme. *Amer. Anthropol.* XLVI, no. 1.
- (17) KRETSCHMER, E. (1925). *Physique and Character* (trans. Sprott). London: Kegan Paul.
- (18) LINDQUIST, E. F. (1940). *Statistical Analysis in Educational Research*. New York: Houghton Mifflin.
- (19) LYND, R. S. & H. M. (1937). *Middletown in Transition*. London: Constable.
- (20) MEAD, M. (1935). *Sex and Temperament in Three Primitive Societies*. London: Routledge.
- (21) MURRAY, H. A. (1938). *Explorations in Personality*. New York: Oxford University Press.
- (22) OLIVER, R. A. C. (1930). The traits of extroverts and introverts. *J. Soc. Psychol.* I, 343-665.
- (23) OSTWALD, W. (1910). *Grosse Männer*, vols. 3 and 4. Leipzig.

- (24) SHERIF, M. & CANTRIL, H. (1947). *The Psychology of Ego-Involvements*. London: Chapman and Hall.
 (25) TERMAN, L. M. & MILES, C. C. (1936). *Sex and Personality*. New York: McGraw-Hill.
 (26) WHEELER, O. A. (1945). *The Adventure of Youth*. London: University of London Press.
 (27) YATES, F. (1946). A review of recent statistical developments in sampling and sampling surveys.
J. R. Statis. Soc. CIX, part 1.

APPENDIX I

SCHOOL..... CLASS.....

BOY or GIRL AGE.....years. DATE OF BIRTH.....

If you have any brothers or sisters, write their *ages* in the next two lines:

OLDER BROTHERS..... OLDER SISTERS.....

YOUNGER BROTHERS..... YOUNGER SISTERS.....

Do any other children *younger* than you, but *not* your brothers or sisters, live in your house all the time?
Underline: YES or NO.

What *kind* of work does your father do?.....

Does your mother go *out* to work? *Underline: YES or NO.*

Underline if any of the following live in your house regularly:

Grandmother/Grandfather/Aunt/Uncle/Lodgers

Is your family: 'Church of England'? or 'Free Church'? or 'Chapel'? or 'Roman Catholic'? or 'Jewish'?
 or 'No Church'? *Underline* whichever one best describes your family.

Put a large $\times$ after each of the following which you have DONE MORE THAN ONCE, and *underline* which-
 ever word describes best how you *felt* about doing that particular thing:

- | | |
|--|-----------------------------|
| 1. Made a toy for a child. | Didn't mind/Liked/Disliked. |
| 2. Looked after younger brothers or sisters or other children. | Didn't mind/Liked/Disliked. |
| 3. Read from books to younger children. | Didn't mind/Liked/Disliked. |
| 4. Dressed a child aged five years or less. | Didn't mind/Liked/Disliked. |
| 5. Fed a baby with a spoon. | Didn't mind/Liked/Disliked. |
| 6. Fed a baby with a bottle. | Didn't mind/Liked/Disliked. |
| 7. Bathed a baby. | Didn't mind/Liked/Disliked. |
| 8. Prepared a <i>cooked</i> meal at home for two or more people. | Didn't mind/Liked/Disliked. |
| 9. Taken baby or young child for a walk. | Didn't mind/Liked/Disliked. |
| 10. Put a baby or young child to bed. | Didn't mind/Liked/Disliked. |

Anything else? If so, write it in here.

APPENDIX II

*Questionnaire on education for parenthood**Form 'A'*

- | | | |
|---|----------|-----------|
| 1. Young people in their 'teens' aren't interested in becoming fathers and mothers. | | |
| Agree | Disagree | Undecided |
| 2. Most parents to-day start their families knowing practically nothing about bringing up children. | | |
| Agree | Disagree | Undecided |
| 3. After marriage is soon enough to learn about bringing up children. | | |
| Agree | Disagree | Undecided |

4. Boys are not interested in the care of children.	Agree	Disagree	Undecided
5. To be a parent needs training just like any other skilled job.	Agree	Disagree	Undecided
6. Looking after young children in the home is a woman's job, and men cannot be expected to take much share in it.	Agree	Disagree	Undecided
7. Practical training in parentcraft should have a place in the work of the secondary school.	Agree	Disagree	Undecided
8. People about to become parents should be provided with some form of instruction about the upbringing of children.	Agree	Disagree	Undecided
9. Fathers are not very important in the care of young children.	Agree	Disagree	Undecided
10. Girls should learn to look after young children.	Agree	Disagree	Undecided
11. Fathers don't take a big enough share in bringing up children.	Agree	Disagree	Undecided
12. Getting some idea of how to be a successful father or mother would be more use in life than are many present school subjects.	Agree	Disagree	Undecided

Form 'B'

1. Young people in their 'teens' are quite interested in becoming fathers and mothers.	Agree	Disagree	Undecided
2. Not many parents these days start their families knowing nothing about bringing up children.	Agree	Disagree	Undecided
3. After marriage is not soon enough to learn about bringing up children.	Agree	Disagree	Undecided
4. Boys are interested in the care of children.	Agree	Disagree	Undecided
5. To be a parent does not require training.	Agree	Disagree	Undecided
6. Men and women should share in the job of looking after children in the home.	Agree	Disagree	Undecided
7. School is not the place for training in parentcraft.	Agree	Disagree	Undecided
8. People about to become parents do not require any form of instruction about the upbringing of children.	Agree	Disagree	Undecided
9. Fathers are quite important in the care of young children.	Agree	Disagree	Undecided

10. Girls do not need to learn how to look after young children.

Agree

Disagree

Undecided

11. Fathers take a big enough share in bringing up children.

Agree

Disagree

Undecided

12. Getting some idea of how to be a successful father or mother would not be any more use in life than are many present school subjects.

Agree

Disagree

Undecided

APPENDIX III

Pressey Collins interests list

Directions. Read over the twenty-five lists of words given below, and draw a line under *everything* you like or are interested in. You may underline as many or as few words as you wish. But be sure to underline everything that you like.

1. fortune-telling boating beaches mountains cinemas.
2. camping tennis football eating milk-bars.
3. Beethoven Faraday Napoleon Raphael Tennyson.
4. kissing flirting pretty girls talkative girls athletic girls.
5. studying dancing day-dreaming walking reading.
6. talking elocution acrobats minstrels smoking.
7. saxophones bands flutes banjos singing.
8. jazz ragtime hymns waltzes solos.
9. good boys handsome boys rough boys rich boys leaders.
10. fashions society clothes banquets games.
11. palmistry pageants card-parties clubs socials.
12. coffee onions mushrooms mustard ice-cream.
13. sleeping teaching clerking loafing typewriting.
14. actors musicians aviators engineers artists.
15. Mowgli Tarzan D'Artagnan Hamlet Galahad.
16. revivals deacons prayer priests church.
17. French drawing English history science.
18. babies sports sailors children animals.
19. college business farming housekeeping travelling.
20. business-men salesmen nurses teachers soldiers.
21. newspapers poems books magazines debates.
22. cards bargains switchbacks machinists electricity.
23. chauffeurs doctors professors bankers grocers.
24. hotels country cities crowds circuses
25. racing seaside fishing arguing joy-riding.

Do not change any of the marks you have already made; leave them just as they are. But now go over each line carefully, and draw a circle round the *one* thing in each line that you like *best*. Be sure you have a circle round one word and only one word in each of the twenty-five lists. Do not miss out any list.

APPENDIX IV

Domestic situations

The situations described below have actually happened. Read each one, and then *write your guess* as to what happened next:

1. Mary got a very poor report on her school work. She took it home and showed it to her father. What did her father do?

2. Amy's mother asked her to go to the shop for her. Amy was reading and did not wish to be bothered. What did her mother do?

3. John and his sister were quarrelling. Their father came into the room. What did he do?

4. Jack got angry with his brother and threw a heavy stone at him. What did their mother do?

5. Ted's father told him to come straight home from school. Ted went to a football match with a chum, and got home at eight o'clock. What happened when he got home?

6. It was a cold wet evening, and there was an hour before bedtime. The children asked their mother, 'What shall we do next?' What did she say?

7. Uncle Bill was coming for the evening. Dick's father and mother wanted to give him a good time. What did they do in the evening?

8. Betty was eight years old. Her parents were going away for a week leaving Betty at home with her older brother and their aunt. Betty threatened to run away if they did not take her with them. What did they do?

Twenty domestic tasks

Here is a list of the jobs which somebody has to do in most families. *Write on each line* the persons whom you think ought to *take a regular share* in doing that particular job:

1. Wash the dishes.

2. Look after the baby.

3. Earn money for the family.

4. Cook the meals.

5. Lay the table.

6. Help the children get ready for school.

7. Go errands to the shops. _____
8. Help the children with their lessons. _____
9. Clean the rugs and carpets. _____
10. Dust the furniture. _____
11. Dry the dishes. _____
12. Clean the house. _____
13. Wash the clothes. _____
14. Serve the meals. _____
15. Open the door when somebody knocks. _____
16. Put the small children to bed. _____
17. Teach the children how to behave. _____
18. Lay and light the fire. _____
19. Wash the floors. _____
20. Look after the children when they get home from school. _____

APPENDIX V

Pity test

Here is a list of things that sometimes make people *feel pity*. After each thing mentioned, underline the words which say *how much* pity you feel:

- | | |
|--|-----------------------------|
| 1. A mouse caught by a cat. | Very much/Much/Little/None. |
| 2. An orphan child. | Very much/Much/Little/None. |
| 3. A bird with a broken wing. | Very much/Much/Little/None. |
| 4. A fish caught on a hook. | Very much/Much/Little/None. |
| 5. An insane (mad) person. | Very much/Much/Little/None. |
| 6. An injured horse. | Very much/Much/Little/None. |
| 7. An underfed child. | Very much/Much/Little/None. |
| 8. An old person with a fatal disease. | Very much/Much/Little/None. |
| 9. A bee drowning. | Very much/Much/Little/None. |
| 10. An old lady who has always wanted children of her own. | Very much/Much/Little/None. |

(Manuscript received 22 September 1950)

THERE IS NO PROBLEM OF MEANING

By GEORGE HUMPHREY

Institute of Experimental Psychology, University of Oxford

I. *The present situation* (p. 238). II. *Psychological difficulties* (p. 238). III. *Perception* (pp. 239–240). IV. *Images* (pp. 240–241). V. *Language* (pp. 241–242). VI. *General ideas* (p. 243). VII. *Thinking* (pp. 243–244). VIII. *Problems for the future* (p. 244). IX. *Summary* (p. 244). *References* (pp. 244–245).

I. THE PRESENT SITUATION

The term 'meaning' continues to be an ambiguous one in psychology. The new Boring, Langfeld & Weld⁽⁴⁾ states, for example, that objects have meaning, disclaims a 'dictionary definition', and gives partial definitions in terms of behaviour and relations. Hebb⁽⁶⁾ defines the 'meaning' of a concept in terms of its 'associations with other concepts'. Murphy⁽¹⁴⁾ defines the term as 'The way in which one interprets a stimulus or a situation in relation to other situations', and for him symbols, to which reference is made, are 'stimuli accepted as *standing for* absent stimuli' (present author's italics). Woodworth's *Experimental Psychology*⁽¹⁹⁾ has ten references to meaning, but no statement about its implications. Writing in 1933, Boring⁽³⁾ refers to Titchener's context theory, which he accepts and amplifies by specifying that 'meaning is response', thus marrying behaviourism and structuralism.

This is, of course, a sample only of the psychological treatments of the term in recent times. Still more numerous and perplexing are the philosophical analyses, for the most part antedating them, and the psychological discussions these have influenced. In 1932 Ogden & Richards in chapter 9 of *The Meaning of Meaning*⁽¹⁶⁾ listed sixteen main definitions which reputable students have favoured, this chapter following one entitled 'The Meaning of Philosophers', which devotes over twenty pages to an almost unadulterated listing of philosophic usages. In his *Signs, Language and Behaviour*⁽¹³⁾, Charles Morris adds a number of other usages, some philosophical and some psychological. It is fair to say that if a layman were to ask a psychologist what he is to understand by the term when used in a psychological context, the proper answer is: 'It depends on which psychologist you are reading.'

This paper proposes to show that the psychological 'problem of meaning' is not a problem at all; that, further, it appears to be a problem only because we have inherited the method of thought which concentrated attention on man and neglected his environment, or, stated otherwise, because the biological aspect of human organisms and their activities has been forgotten. With any philosophical problems it has nothing to do.

II. PSYCHOLOGICAL DIFFICULTIES

The principal points where the term meaning causes difficulty in psychological theory are perception, images, language (normal and pathological), general ideas, thinking and possibly dreams and other psychoanalytic interpretations.

III. PERCEPTION

Starting with perception, we may take Newman's admirable statement in a passage already referred to (4): 'Perception is the experience of objects and events which are here now.' It is, however, unfortunate that later on he adds: 'The fourth important characteristic of objects is that they have *meaning*. Every common object in your view is thoroughly familiar to you. You not only see it but you also recognize it as being such and such an object.' Meaning is here more or less equated (psychologically) with recognition or familiarity. In the same page and column it is equated primarily with *activity* ('what an object means to us is first of all what we do about it'), and secondarily, in the next column, with *relation* to other objects.

This account is confusing. To equate meaning and recognition is surely an explanation *ignoti per ignotius*; and to amplify recognition into *organic reaction to an object*, and *relations between objects* seems to confound confusion worse. It is true that perceived objects are generally recognized, that we react to them, and that we perceive them as related. But why use the word meaning at all, especially in a new sense, 'not the same as a dictionary definition'?

The answer probably lies in history. Psychologists have been accustomed to speak of perception in terms of meaning. Thus, to go back to Helmholtz, writing in 1867, we find that sensations (i.e. primary sense data, if one may use that fighting phrase) are symbols for the relations in the external world. They 'enable us to form conceptions as to the existence, form, and position of external objects. These ideas are called *visual perceptions*.' Perceptions of external objects are, further, of the nature of ideas. 'Our ideas of things *cannot* be anything but symbols, natural signs for things' (7). We must make '[unconscious] judgements, in which our sensations... are connected with the existence of an external cause'. Now the word meaning is not expressly used, but it is definitely implied by the words, symbols and natural signs. The main point is that there is the thing and the symbol for the thing, which 'belong to two different worlds'.

A similar doctrine seems to have been held, at least at one time, by William James (10), who, in spite of his insistence that sensation has an object, still speaks of a harmony evolving between the outer and inner life. And he does contrast the 'childish perception of stains of colour merely as such' with the 'consciousness of what they mean'. Our 'inner faculties are *adapted*... to the features of the world'.

But enough has been said to illustrate the fact, if illustration is needed, that the concept of perception has been bound up with the concept of symbol and of meaning, the latter word often being italicized in the discussions (Angell (2), James (10), Murphy (15)). It is fair to say that examination of treatises dealing with perception reveals great confusion, the term meaning being generally found or implied somewhere in the account, but being used in different senses by different writers. Actually it is not necessary at all, if one lives up to the psychological realism which most psychologists now imply in their writings. As Murphy puts it: 'psychology is the science which treats of the interrelations between the environment and the responses of the organism' (14), a definition which is perhaps a little broad, but clarifies the point at issue. One of these interrelations, as Murphy remarks and as Boring would agree, is the direct perception by the organism of the environment. There is no point in bringing into the picture a *tertium quid* such as an idea or a sensation, nor need we use in our primary statement any such term as sense datum.

We perceive the world directly, and when, as Newman says, we wish to focus attention on the contribution of our sense organs, we speak of *sensation*. As soon as we forget that *to perceive* is a transitive verb, and thus separate it from the *environment perceived*, we find ourselves in the world of primary sensation, presentations, ideas, the 'two worlds', and so on, from which we can never escape except by a pre-established harmony, an act of 'reference', by the assumption of 'intentional inexistence', or other postulated mechanism. Or we have to speak in terms of symbols, which have the outside world as their 'meaning'. The 'problem of meaning' in perception is an artificial one made necessary by putting a *tertium quid* between the process of perceiving and the environment perceived, and made even more difficult by the invention of the word 'percept'.

IV. IMAGES

Here the 'problem of meaning' seems to come out with special clearness. 'I have an image of a boy in a blue suit' seems to be a reputable psychological statement. Then immediately one is justified in asking: 'What is the meaning of the image?', and the answer seems to be, of course, 'the boy'. This is surely the 'superficial, common-sense meaning' (Titchener⁽¹⁸⁾), which may have to be amplified on further consideration, introspective or otherwise. But there are difficulties. Titchener's context theory, which is related to James's 'fringe of consciousness', demanded that the meaning of an image is another image, i.e. another mental process. For him, it takes two mental events to make a meaning, and we are once more in the 'two different worlds' theory, according to which we can never get over to the actual environmental 'boy in blue' (as contrasted with 'mental' process or state) demanded by psychological realism and psychological common sense alike.

How, once again, to escape over the charmed gap? It is possible of course to bring in a *deus ex machina*, and indeed a whole pantheon, or rather theomachy, has actually been created. Thus it is possible to speak of the mental function *sui generis* by which we refer to the boy (reference or referential function), and this has the advantage that it may be used in a perfectly neutral sense; or we may, as did Moore⁽¹²⁾, postulate a fresh (in his case fourth) category of mental life, namely, meaning; or we may speak of relation, or, with Titchener and Boring, context. In the latter cases, the question must be asked, what sort of relation or context? Boring has ingeniously answered this question, as Watson did before him, by saying that meaning is given by reaction ('the dog means what it does', said Watson on a notable occasion). It does not seem that Boring's amplification of Titchener's statement is satisfactory. I cannot, for instance, believe that the thrashing I am about to administer to the boy can be said to be the 'meaning' of the 'visual image'.

To answer the question how the image can mean the boy one must go back to the primary psychological datum. The organism is doing something, namely, *imagining* the boy. One does not make an image which, one realizes through a separate mental process, is the image of the boy, any more than one first makes a presentation, idea, or percept by means of which one thereafter perceives him. The term image is clearly superfluous as a psychological description of this primary psychological process and the problem of the meaning of the image has evaporated. Dawes Hicks⁽⁸⁾ prefers to use the term 'imaging', because 'imagination' includes so many different processes. But on the whole it seems better to call the process one of imagining, and to use 'imaging' or even the dangerous 'image' to describe that aspect of it which may be classified under one or other of the sensory modalities.

The objection immediately occurs that we cannot here be doing something that is affected by the nature of the environment, since the particular piece of the environment a question is not present, and in fact may not now be in existence at all. The answer is that we do exactly this thing, though we generally call it memory or learning. There is certainly a problem of memory. It is not the 'problem of meaning'. Neither here nor elsewhere is there a specific 'problem of meaning'.

It should be added that this paper does not maintain that the term 'image' should be abolished, but that for a psychological description of primary organic activity it is unnecessary. The term is profitably used with caution to describe a specific organization involved in the process of imagining. Caution is necessary, since here, as with the correlative term 'percept', it is so easy to slip into forgetting that the organism is *doing* something in which the organism's environment plays a part. Once again, the problem of the meaning of an image is an artificial one brought about by separating the organism from its environment.

V. LANGUAGE

Here the problem of meaning seems to emerge real and lifelike. Two people are talking. One is making certain sounds, and we say he is speaking German. A bystander asks: 'What does he mean?' and is told 'He is saying that the car is outside'. Here surely is a problem of meaning, in that one must ask how the noises made by the speaker come to mean anything about the car. It has already been seen that no help can come from the 'referential function', context, relation, fringe of consciousness, intentional inexistence, or even 'intention', the scholastic term. These are *ad hoc* terms, which do but show our ignorance. The primary psychological fact is that *A* is reacting to a total situation comprising (i) *B*, the hearer, together with (ii) the car which is outside. *B* answers, perhaps, 'I will get my coat'. In his turn *B* is reacting to *A* in front of him, and also to the car outside and to his coat. How can noises produced by *A* bring a reaction to the car outside and to the coat? This is, indeed, a problem, *but it is not a problem of meaning*. It is a problem of learning, sometimes called the problem of reduced cues, sometimes the problem of substitute activity (wrongly at least in the case of the conditioned reflex, and probably in any other sense), and, in the opinion of the writer, really a problem of four-dimensional integrative reaction⁽⁹⁾. Whatever the nature of learning may be, and whatever its problems, there is no special problem of linguistic learning, and still less a problem of linguistic meaning.

To leave the statement at this point would be disingenuous. Not only does *B* respond by saying 'I will get my coat', but he also understands *A*'s meaning. This involves the further question of thinking, which will be considered later. It should also be pointed out that both *A* and *B* must know the language, so that the 'meaning' is common to both, or there will be 'misunderstanding'. It is, however, not difficult to see how *A* and *B* may both have learned approximately the same thing, although the one hears and the other speaks. Modern learning theory has developed far enough to explain roughly what has happened. In order that learning may take place, the organism is ordinarily confronted with approximately the same repeated situation.¹ The learned response, in its broadest sense, gradually develops. The set of physically discrete situations which have faced the organism may be called collectively the *learning series*; after learning has taken place they are no longer discrete,

¹ The single presentational learning stressed by Gestalt psychologists is taken as a limiting case.

as far as the organism is concerned, since they have now been related by the organism in the process of learning. Thus twenty successive presentations of the maze are related by the rat, the last 'run' being different because of the preceding events. Twenty golf 'lessons' are related by the golfer, and succeeding performances are different because of them. This formation of a fresh response is part of the total process which the writer prefers to call 'four-dimensional integration'. But Lashley and many others have shown that the response which is the end-product of learning need not necessarily be identical with any of the responses that took place during learning. An animal which has been operated upon so that it cannot walk in the ordinary way may roll or lunge through the maze; another which has learned to run through may swim through; another may take a three-quarter right turn in order to go to the familiar left, and so on. These well-known cases are all organizations of response made at the end of a learning series to meet the situation immediately confronting the animal. The same thing happens with human beings, whose learned responses are of course much more flexible. Sara Fisher⁽⁵⁾ has shown, for example, that the end-product of learning may include an 'image' (a visual organization which itself developed as learning progressed), or it may be a word repeated, a *name*, in this case a nonsense syllable, such as *Deral*; or it may be a 'verbal description' of the 'nonsense figure' in question. The verbal description is also a set of learned responses. It is a verbal organization effected when the organism is faced by a complex situation, including a 'social stimulus', with roots reaching into the past through a great number of learning series. It is, in fact, part of a four-dimensional integration which involves the present situation, social and otherwise. The word 'meaning' is not necessary in order to describe the psychological events that have occurred either in the case of the 'image' or of the 'verbal description'.

Thus in this experiment subjects learned to say the word *Deral*, and to imagine 'a *Deral*'. The latter process included the formation of a visual organization called an 'image', the stages of this part of the learning being described in detail. At the same time subjects had learned to comprehend and to react to the *spoken* word *Deral* when it formed part of a spoken context. All these learned processes came as the result of the same learning series, through the four-dimensional integration effected by the human organism. There were in Fisher's experiment five subjects, all of whom had learned, through approximately the same learning series, to say, to comprehend and to react to the *spoken or written word Deral*.

The above argument does not, of course, depend on the validity of the concept of four-dimensional integration. At this point, however, it may be objected that the notion of meaning, our ignorance of which is at least clear, has been explained away by introducing a new term concerning the implications of which we know nothing. It may then be in place to say that the concept was used, though without the name, by Sir Charles Sherrington forty years ago, when he employed the word 'integration' with reference to the work of the distance receptors, the field of operation of whose arcs brings 'higher feats of associative memory'⁽¹⁷⁾. If it be objected that this is three- and not four-dimensional integration, the answer is that memory is of time gone by, and thus a four-dimensional process is necessarily involved. And animals have never, and human beings have only for three hundred of their million years, known that Newton divided the flux of environmental events into three static dimensions of space together with a unique extra dimension of time. All integration is necessarily four-dimensional.

VI. GENERAL IDEAS

Here again the term 'meaning' seems almost necessary. How is a 'general idea' possible? The problem was hotly debated by the philosophers, through whose analysis we can now see the question more clearly. That generic images exist which are related to a class of objects rather than particular objects there seems now no doubt. It has been shown also that *qua* images they need not have any special characteristics. Thus it is usual to say that the image may be particular but the reference general, a usage with which the writer was long content. Binet stated that an image may stand for anything. William James has made the point with admirable lucidity. 'The geometer with his one definite figure before him knows perfectly that his thoughts apply to countless other figures as well, and that although he *sees* lines of a certain special bigness, direction, colour, etc., he *means* not one of these details' (11).

In these two statements we have the term meaning, together with 'stands for', a metaphor which, like 'vehicle of meaning', is often used as though it were an explanation. After the previous discussion the main lines of criticism are clear. To take Sara Fisher's study as an example, what happened was this. One (alternative) response made possible by the process of learning was that the subject *imagined* the nonsense figure, e.g. a *Deral*. When asked to *describe a deral*, this process of imagination was called into play, the image being that aspect of it, that organization entering into the total response, which can be classified under the visual modality. According to Ach's⁽¹⁾ theory of convergence, the image or idea in these cases corresponds to an average of the actual members of the class. Since the image is one of the mechanisms acquired in the process of learning to respond to an individual object as a member of a class, there is no reason to ask its meaning, nor how it can *stand for* a class, nor how it can be the '*vehicle*' of this, that or the other meaning. The total process of imagining the *Deral* is learned when the organism has been faced with the learning series consisting of so many *Deral* figures. It is, to repeat, but one of the responses thus learned. The image itself is an organization, itself a more or less completed end-product, contained within the total end-product of learning. As long as the process of learning, its integrative nature, and its environmental correlative, are kept in mind, the phrase 'the meaning of the general image' is grotesquely superfluous.

VII. THINKING

At first sight this seems the ultimate problem. How can a train of thought which, whatever it is, is obviously not 'what I am thinking about', come to *mean* 'what I am thinking about'? We have seen that we perceive the object directly, and that as soon as psychological theory separates the process of perceiving and the object of perception, the problem of meaning rears its head. A similar result comes when the same separation is made between the organism and the object imagined, the object spoken of and to, or the group of objects to which generalized response is made. Psychology studies organisms doing things to their environment. Can we *think* a thing, as we perceive, imagine, or conceive it? According to the Würzburg psychologists and the other advocates of imageless thought, that is exactly what we do. They maintained that we may think of an object without the intermediary of perceptions, presentations, images, or any other processes which can be classified under one or other of the sensory modalities. On the theoretical side their opponent

was Wundt, on the experimental Titchener. Titchener claimed in effect that we think through images, which were the vehicle of meaning, and that meaning was given in consciousness by another image. The urgency of Titchener's problem, and in the writer's opinion he never solved it, came from the fact that he introduced the image between thought and its object. Once an image, then follows necessarily the problem of meaning, and for Titchener's system, necessarily another image, if the writer's interpretation of him is correct. Of course, Külpe's school was not the only advocate, experimental or otherwise, of imageless thought. Many psychologists of the highest rank, including Binet and Woodworth, as well as many philosophers, have come to much the same conclusion. To the writer, after reading the Würzburg work, the theoretical criticisms of Wundt, and the experimental ones of Titchener's laboratory, there is no doubt on which side lies truth. To debate the question of imageless thought would not be in place in this paper. It can only be said that if the general thesis of the Würzburgers is here accepted, and in the opinion of the writer it must be accepted on other grounds as well as those brought forward by the Würzburgers, then the problem of the meaning of thought evaporates, although the Würzburgers did not see this at the time.

VIII. PROBLEMS FOR THE FUTURE

It seems probable that in relation to dreams the term meaning will prove likewise to be founded on an artificial treatment of the facts, but the writer is not yet prepared to make a definite statement on this or the general psycho-analytic use of the term.

The foregoing statement raises its own problems. Of these the chief, perhaps, lies in the statement that, as well as perceiving an object directly, we may also imagine and think it directly. Are there then three distinct processes? Three mental operations, psychological 'things' we do to the surrounding world? Probably not. The relation between perception, imagination and thinking has long been debated; it may turn out that they form a continuum, along which something like the internal and external 'forces' of the Gestalt psychologists play a greater or lesser relative part. This also must at present be considered a problem for the future. At the present time it looks to be not insoluble.

IX. SUMMARY

It is maintained that by the artificial separation of psychological processes from the environment in which they function an unnecessary 'problem of meaning' has been created.

REFERENCES

- (1) ACH, N. (1921). *Ueber die Begriffsbildung*. Bamberg.
- (2) ANGELL, J. R. (1908). *Psychology*, p. 247. New York.
- (3) BORING, E. G. (1933). *The Physical Dimensions of Consciousness*, p. 223. New York and London.
- (4) BORING, E. G., LANGFELD, H. S. & WELD, H. P. (1948). *Foundations of Psychology*, pp. 230, 217, 230. New York and London.
- (5) FISHER, S. C. (1916). The process of generalizing abstraction, etc. *Psychol. Monogr.* xxi, 90.
- (6) HEBB, D. O. (1949). *The Organization of Behaviour*, p. 130. New York and London.
- (7) HELMHOLTZ, H. VON (1925). *Treatise on Physiological Optics*. Vol. III, *The Perceptions of Vision*, pp. 18, 1, 19, 27. The Optical Society of America. (Original written in 1867.)
- (8) HICKS, G. DAWES (1924). On the nature of images. *Brit. J. Psychol.* xv, 121-48.
- (9) HUMPHREY, G. (1933). *The Nature of Learning*, chapter 5. London and New York.

- 10) JAMES, WILLIAM (1892). *Textbook of Psychology* ('an abridgement of my larger work'), pp. 4, 344, 3. Macmillan.
- 11) JAMES, WILLIAM (1907). *Principles of Psychology*, I, 472. London.
- 12) MOORE, T. V. (1915). The temporal relations of meaning and imagery. *Psychol. Rev.* XXII, 177-225.
- 13) MORRIS, C. (1946). *Signs, Language and Behaviour*. New York.
- 14) MURPHY, G. (1935). *A Briefer General Psychology*, pp. 564, 11. New York and London.
- 15) MURPHY, G. (1929). *An Historical Introduction to Modern Psychology*, p. 190. London and New York.
- 16) OGDEN, C. K. & RICHARDS, I. A. (1932). *The Meaning of Meaning*, p. 186. London and New York.
- 17) SHERRINGTON, SIR C. S. (1908). *The Integrative Action of the Nervous System*, p. 349. London.
- 18) TITCHENER, E. G. (1912). Description *v.* statement of meaning. *Amer. J. Psychol.* XXIII, 165-82.
- 19) WOODWORTH, R. S. (1938). *Experimental Psychology*, index. New York.

(Manuscript received 22 February 1951)

TOWARD THE REINSTATEMENT OF THE CONCEPT OF THE SELF

BY LEON LITWINSKI, *Brussels*

I. *Introduction* (p. 246). II. *Possession and seeking as features of the self* (pp. 246-247). III. *Expectation* (pp. 247-248). IV. *Conclusion* (pp. 248-249). V. *References* (p. 249).

I. INTRODUCTION

Recognition that the search for the self is a significant scientific endeavour seems to be increasing. The thesis chosen by E. R. Hilgard⁽³⁾ for his address as president of the American Psychological Association, viz. that the adoption of the concept of the self is essential for understanding defence mechanisms, suggests that this concept and its references are coming once more into the forefront among the preoccupations of contemporary psychologists.

Some twelve years ago, G. W. Allport⁽¹⁾ emphasized that the central position of self is implied in all states of consciousness; such concepts as function, adaptation, use and adjustment are of no significance without reference to the person. In this connexion, Allport pointed out some plausible arguments through which personalistic psychologists like W. Stern⁽¹⁷⁾ and self-psychologists state their case for the reconstruction of psychology.

At present, the stimulating tendency to restate in new terms some ancient ways of thinking may be observed in the handling of traditional psychological problems by Snygg & Combs⁽¹⁴⁾. The thesis of these authors is that behaviour is not determined by the situation, but by the way one perceives the situation, i.e. by the present phenomenal field of those who behave. We are far from the time of Taine, who put excessive emphasis on the environment.

On the other hand, if the old adage is true, and certainly it is, that 'everyman has three characters: that which he has, that which he thinks he has, and that which others think he has', we can more easily understand why Hilgard⁽³⁾, with a view to getting a better knowledge of the self, rejects a purely introspective approach to the search for the self and seeks a more inclusive self-concept, whereby 'the person comes to see himself as an informed other person sees him'. The 'self', considered by this objective attitude, Hilgard calls the 'inferred self'.

II. POSSESSION AND SEEKING AS FEATURES OF THE SELF

The protagonists of the 'phenomenal self' as well as those of the 'inferred self' seem to share with G. W. Allport⁽¹⁾ and G. Murphy⁽¹³⁾ and others like them, the view that one of the fundamental characteristics of self-awareness is 'an evaluative or judging attitude toward the self, in which the self is regarded as an object of importance, and preferably of worth' ⁽³⁾. A similar view has been expressed in different terms by the author of the present paper ^(9, 11). But all this is only a half truth. The self can be viewed not only as evaluating or evaluated, i.e. as the subject evaluating or the object evaluated, but as a gratified seeker or possessor as well.

Our self-esteem rises in our own eyes with the gratification of our seekings, i.e. with

the acquisition of things, social links, ideas. It arises with the acquisition of a house, a motor-car, a social function, a literary work, with the discovery of a thing or an idea.

The question which I wish to consider, and which I have raised in the past, is whether, side by side with the self mechanized for its defence and studied more particularly in its negative aspects by Freud, Adler, Jones and many others, the 'self-psychologists' should not put at least equal stress on the self considered in its positive aspects. The point is whether in insisting on the self that feels guilty and is anxious, whether in studying the ways in which we deceive ourselves by denial of impulses, of traits, or of memories or by disguise in assigning our own undesirable traits to other people, in order to maintain or to restore self-esteem, we do not lose sight, far more than we should, of the self that is flattered, that glorifies itself openly and frankly, that is proud of its acquisitions or achievements and that consequently likes itself.

Despite the difficulties in introspective approaches to the self, Hilgard⁽³⁾ finds that our self-concept needs to include some information based on private experience to enrich the concept of the self and to make it square with all that is known about human motivation.

The private experience of the self expresses itself to a great extent in acquisitive tendencies¹, i.e. in the prospective links constituted by the evaluating self, but nevertheless the self does not identify itself purely and simply with these links. The outer expression may understate or overstate the level of the self. William James⁽⁵⁾ was right when he said that 'in the widest sense of the word, self envelops everything that a man can call his'. This statement has to be developed. As I pointed out elsewhere⁽¹¹⁾, what has been effectively selected does not necessarily correspond exactly to what is valued. One is often forced to acquire what he finds. *Lucri quidquid est, id domum trahere oportet*, said Plotinus. Knowledge of the things, ties or ideas acquired by the person is highly helpful in such interpretations.

III. EXPECTATION

By the end of the eighteenth century, Bentham⁽²⁾ stressed, in connexion with his analysis of the different titles to property, that a continuity exists between the person and its objects because man is capable of prospective thinking. According to Bentham, 'property is nothing more than the basis of a certain expectation; namely, the expectation of deriving hereafter certain advantages from a thing (which we are said to possess) in consequence of the relation in which we stand towards it'.

If neither lawyers nor psychologists paid attention to Bentham's contribution with regard to the idea of expectation, it must at least be said to the credit of psychology that a century after the publication of the French edition of Bentham's treatise, Clara M. Hitchcock⁽⁴⁾ published a monograph entitled *The Psychology of Expectation*. In this study the author says that unlike memory and other mental processes, expectation has received scant mention in the writings of early philosophers, while modern psychologists are content with a brief account of its workings. It is classed by them variously: sometimes with belief, sometimes with feelings or emotions, and, again, with the cognitive processes.

¹ Generally speaking, psychologists classify acquisition side by side with other basic needs like affiliation, retention, recognition, seclusion, etc. To my mind, acquisition may be conceived differently, i.e. as a fundamental tendency of the person to create prospective links with various activities in order to gratify mainly the need of security taken in its most comprehensive sense. These links express tastes, preferences, affinities, style of the self. Thus affiliation is a kind of acquisition.

Hitchcock places it in the last class, primarily because its reference is distinctly objective, pointing to that which has definite time and space relations beyond the individual mind, whereas emotion pertains to the subjective side of an experience.

Although Hitchcock does not refer anywhere to Bentham's theory, she admits in her study that there are multitudes of commonplace anticipations that serve their purpose of preparation in the routine of life and pass unnoticed. As a matter of fact they serve much more than the above purpose; their object is to maintain security for the preservation and development of the organism that achieves forethought.

If we turn now to more recent writings on expectation, E. L. Thorndike⁽¹⁸⁾ admits that what has come to be known as S-R psychology has paid little attention to expectations. After discussing various views on the interpretation of expectation, Thorndike suggests some observations and experiments to get a satisfactory account of primitive forms of knowledge and how they operate.

According to Tolman⁽¹⁹⁾, expectancies, called previously expectations, need not be, or be accompanied by, any conscious ideas, images, hopes, fears, impulses or the like. He maintains that an intelligent functioning, i.e. the successful reaching of a goal, is an expectancy on the part of the subject.

Whatever may be the future of these attempts to interpret animal behaviour in terms of expectation, they do not seem, at least at present, to have a direct bearing on our subject.

There are various forms of expectation ranging from perception to highly conceptual thought. The particular form of expectation which I have termed 'relaxed expectation'⁽¹⁰⁾ is a function of security and may correspond to what Ross Stagner⁽¹⁵⁾ calls a positive valence in expectancy. According to this author, the function of expectancy valence is itself a form of anticipation. If a person anticipates hunger, the memory of discomfort and the expectation of future gratification may cause food objects to have positive valence even in the absence of a physiological need for food.

IV. CONCLUSION

The concept of the self and its references raises a certain number of fundamental questions for contemporary psychology. The self seeks and evaluates. It evaluates things with reference to present or future needs and wishes. As long as the distinction between the self and the environment is badly defined, the concept of a thing's belonging to one's own self is not made clear. According to W. Stern⁽¹⁶⁾ it does not appear for the child until the second or third year. The self feels, knows and acts as seeker or possessor. It likes or dislikes, it is proud or ashamed, stable or fluctuating. Within varying situations and action patterns it is aware of its identity as well as of its continuity.

Psychologists seem to be differently impressed by the above distinctions. James conceived the ego and its references as the unifying factor running through all our experiences, as the stream of thought and an empirical aggregate of things objectively known. For Kurt Lewin⁽⁶⁾ the unity of consciousness is not identical with the unity of the whole region of psychical forms and processes, of tense and not-tense psychical systems, the totality of which may be designated as the 'mind'. A number of facts justify the view that a special region, within the psychical totality, must be defined as the self in the narrower sense. The distinction of relatively separate psychical systems leaves open various possibilities for the question of the unity and homogeneity of the self.

Acquisitiveness should not be considered as a self seeking tendency directed exclusively toward material things, but toward social and spiritual constituents as well. Further the definition of personality as an inner system of expectations and desires should be accepted as having numerous advantages. As Ross Stagner⁽¹⁵⁾ says, such a definition is less confusing than a definition in stimulus terms. He describes Allport's definition of personality as 'a good judgement'. As Bentham said, and as I emphasized in a paper published in 1913⁽⁷⁾, expectation and security are intimately connected. The security-insecurity dimension is a fundamental aspect of the self-image⁽¹⁴⁾.

It seems that to the phenomenal field of Snygg and Combs, Koffka's behavioural field, Lewin's life space of the individual, Frank's private world, considered as frames of reference for psychology, one can add what I should call the *expectation realm*, which allows us to discover all these things that we call *our* objects, persons, ideas, etc.

As I emphasized elsewhere⁽⁸⁾, science progresses either like a spiral or a pendulum. A backward reference may sometimes be interesting and useful. Thus McDougall⁽¹²⁾ has pointed out that dynamic theories of psychology reveal strikingly Aristotelian concepts. One of the objects of this paper has been to show the interest of Bentham's principle of *reasonable expectation* from the point of view of a contemporary psychology of the self.

REFERENCES

- (1) ALLPORT, G. W. (1949). *Personality, A Psychological Interpretation*, p. 550. London: Constable Co.
- (2) BENTHAM, JEREMY (1914, 1931). *Traité de Législation Civile et Pénale*. Published in French, in 1802, in Paris by E. Dumont, a Swiss; translated later into English, *Theory of Legislation*, Bowring edition, 1840, Ch. VIII and Ch. XVII. Atkinson edition, London: Noon; Ogden editions, London: Kegan Paul.
- (3) HILGARD, E. R. (1949). Address of the President of the A.P.A. at Denver, 6 September 1949. *Amer. Psychol.* Sept., pp. 377-8.
- (4) HITCHCOCK, C. M. (1903). The psychology of expectation. *Monogr. Suppl. to the Psychol. Rev.* no. 20 (vol. v, no. 3). New York.
- (5) JAMES, WILLIAM (1890). *The Principles of Psychology*, p. 228. New York: Henry Holt Co.
- (6) LEWIN, KURT (1935). *A dynamic Theory of Personality*, p. 56. New York: McGraw-Hill.
- (7) LITWINSKI, L. (1913) Qu'est-ce la propriété? *Rev. Int. Sociol.* Paris: Alcan.
- (8) LITWINSKI, L. (1943). *Les aspects sthéniques et asthéniques dans les conduites et les sentiments complexes*, p. 2. Coimbra, Lisbon: A. M. Pereira.
- (9) LITWINSKI, L. (1947). The psychology of 'mine'. *Philosophy*, Nov. London: Macmillan Co.
- (10) LITWINSKI, L. (1947). Une nouvelle théorie psychologique du lien possésoire: l'attente dans la détente. *Journal de Psychol. norm. path.*, Oct.-Dec. Paris: Presse Univ.
- (11) LITWINSKI, L. (1949). La psychologie du 'mien'. Rapport présenté au XIIe Congrès International de Psychologie, Edinbourg, Juillet 1948. *Acta Psychol.*, July, pp. 194, 197. Amsterdam: North Holland Publishing Co.
- (12) McDOUGALL, W. (1921). *Dynamic Principles in Recent Psychology*. In A. L. Robinson, W. McDougall, a bibliography. Durham: Duke University Press.
- (13) MURPHY, G. (1947). *Personality: A Biosocial Approach to Origin and Structure*. New York: Harper Brothers.
- (14) SNYGG & COMBS (1949). *Individual Behavior, A New Frame of Reference for Psychology*, p. 184. New York: Harper Brothers.
- (15) STAGNER, ROSS (1948). *Psychology of Personality*, pp. 302-3. New York: McGraw Hill.
- (16) STERN, W. (1910). *Erinnerung, Aussage und Lüge in der ersten Kindheit*. Leipzig.
- (17) STERN, W. (1938). *General Psychology from the Personalistic Standpoint*. Macmillan Co.
- (18) THORNDIKE, E. L. (1946). Expectation. *Psychol. Rev.*, Sept., pp. 277-81.
- (19) TOLMAN, E. L. (1945). *Science*, pp. 101, 160-6.

(Manuscript received 27 July 1950)

ON SOME DIFFICULTIES ENCOUNTERED IN THE USE OF FACTORIAL DESIGNS AND ANALYSIS OF VARIANCE WITH PSYCHOLOGICAL EXPERIMENTS

By B. BABINGTON SMITH

I. *Introduction* (pp. 250–251). II. *Experimental procedure used* (pp. 251–253). III. *Inter-personal variation* (pp. 253–261). IV. *Intra-personal variation* (pp. 261–262). V. *Qualitative observations* (pp. 262–265). VI. *Discussion* (pp. 265–266). VII. *Summary* (p. 266). *Appendix* (p. 267). *References* (p. 268).

I. INTRODUCTION

In agricultural and biological experiments where plants or animals have been subjected to a series of treatments, the *order* of treatments has been found to be of importance (e.g. a treatment may permanently injure a plant or animal or plot of ground). Order is more generally important in psychological work. I comment here on results in experiments in which a series of responses are called for. These remarks do not apply to experiments in which one response only is called for from each of a number of subjects or groups.

In the first flush of enthusiasm, analysis of variance has been used very freely and apparently with little regard to possible snags. For example, in a recent text-book (*Statistics in School*, by Sumner⁽⁹⁾) four backward readers, Tom, Dick, Harry and George, undergo four methods of training in spelling, the allocation of 'treatments' to boys being arranged to give a Latin square. To arrange such a scheme and to interpret it by analysis of variance, without a hint that previous training in method A might affect subsequent training by method B or C or D, or, indeed, any subsequent training, is to overlook one of the characteristic features of human beings. Edwards⁽³⁾, in his recent book, includes a chapter on experimental designs involving repeated measurements of the same subjects, but so far as I can see does not raise the issues discussed in this paper. I have myself quite deliberately used the method with extreme forms of data; but I have done so with my eyes open to this extent, that I was seeking for the limits beyond which results given by the method ceased to be sensible or reliable.

The field of application of analysis of variance has steadily grown by exploration; first to cover various special distributions, such as the Poisson, then to variates which by relatively simple transformation, like the square root or logarithmic, can be induced to fall within the scope of normality. None of these methods, however, seriously touches the difficulty which seems to me inherent in all psychological experiments, namely, that the human subject (like the elephant) doesn't forget, but remembers stimuli or responses over relatively long periods. One way of looking at this is to say that all results obtained in a series of responses from one subject must be treated as time series with all the consequent complications. Cochran's⁽²⁾ work on rotational experiments is relevant in this connexion so also is J. E. Walsh's⁽¹⁰⁾ article on the effects of intra-class correlation on significance tests. (It is perhaps reassuring, for present purposes, that Walsh finds the variance ratio test is likely to be less affected by intra-class correlations than χ^2 or t . The risks of the t test are very striking.) On the psychological side there have been relatively few warnings though it has long been a practice to avoid marked contrasts in a series of stimuli involving

different values of some variable, as in paired comparison and the constant method, and a good deal has been said about the effect of expectation (which is in part the outcome of past responses) on a subject's response.

The data in this paper show features which in themselves seem to suggest a warning against direct acceptance of results obtained by means of analysis of variance. That is, where the successive responses are interrelated, there is a risk that the mean responses to the different factorial levels (or, in other words, the 'main effects') are not independent. This is not the same as saying that there is an interaction. It means that the level of response to stimuli of type A, in isolation, is not the same as the response to A's when they are presented, according to some factorial design, in conjunction with B's, or B's and C's, and so on. It is rather as if, in an agricultural experiment, the response to treatments in the various plots in a block was dependent on which treatments were used, not only on the plots themselves but on their neighbours as well. Where we find this phenomenon, i.e. that the level of response to A is in part dependent on whether other levels of the same variate are used, then it is clear that we must be very careful about the interpretation of the main effects. Psychologically, one way of describing the situation, where the main effects are interrelated, is to say that the stimulus situation becomes organized, or takes on a pattern to which response is made. There are various ways for trying to express such an idea; in some ways the situation may be thought of in 'gestalt' terms, in Koffka's⁽⁷⁾ words: 'a gestalt is therefore a product of organization, organization the process which leads to a gestalt'. It is relevant at this point that Katz⁽⁶⁾ has recently published papers on 'gestalt' in mental work.

Beyond this question of interrelated main effects, I also present evidence about the effect of randomization of types of stimuli upon residual or 'error' variance; it is somewhat obscure, but I have some qualitative observations to make.

II. EXPERIMENTAL PROCEDURE USED

In 1934-5, I was very anxious to apply analysis of variance to standard psychological experiments and to design fresh experiments so that the method could be the more effectively applied. I expected that estimates of performance would thus be obtained, free from the usual objection to repeated measures or responses, namely, that subjects adapt themselves to situations and learn by repetition. In this year, and in 1935 and 1936, variants of the 'Muller-Lyer illusion' were used in the psychological laboratories in St Andrews and Dundee as a class experiment, with the usual arrangement such as is set out in Guilford's⁽⁵⁾ *Psychometric Methods*. I treated the results by the method of analysis of variance and in 1936, 1937 and 1938, I had the opportunity of redesigning the experiment and using factorial designs. This, as I have said above, I did expecting to obtain more precise estimates of main effects. An extract from the manual of instruction used in 1936 and 1937 is given as an Appendix.

As the experiment proceeded I began to feel that the randomization involved was helping to retard learning and adaptation. This development promised to be very advantageous because, if confirmed, it would mean that not only was this a way to obtaining more precise estimates of main effects by increasing the number of determinations on which they were based, but also that one might expect still further improvement of estimate by cutting down practice effects. Even the very marked practice effects so

Table 1

Particulars of experimental procedure

(i) Place	(ii) Year	(iii) No. of subjects	(iv) Length of standard line	(v) Types of setting	(vi) Factor pattern	(vii) Replica- tion	(viii) Total no. of readings	(ix) Type of arrangement
Scotland	1934	22	75 mm.	C; R, L; I, O	2×2	20	80	20 each RI, RO, LI, LO in that order; systematic
St Andrews and Dundee	1935	14	75 mm.	A, B, C; R, L; I, O	$3 \times 2 \times 2$	20	240	20 each as above for A, C, B in that order; systematic
	1936	18	75 mm.	A, B, C; R, L; I, O	$3 \times 2 \times 2$	6	72	6 blocks of 12; each combination appears once in each column, once in each pair of rows
	1937 1938	18 15	75 mm. 75 mm.	A, B, C; R, L; I, O A, B, C; R, L; I, O	$3 \times 2 \times 2$ $3 \times 2 \times 2$	6 4	72 48	6 blocks of 12, as in 1936 4 blocks of 12, once in each column, once every three rows
Oxford	1946	7	Convenient lengths measured subsequently	A, B, C; R, L; I, O	$3 \times 2 \times 2$	4	48	4 blocks of 12, as in 1938
	1947	8	50 or 75 mm.	A, B; R, L; I, O	$2 \times 2 \times 2$	8	64	8 blocks of 8, Latin square
	1948(I)	6	Convenient lengths measured subsequently	A, C; R, L; I, O	$2 \times 2 \times 2$	8	64	8 blocks of 8, Latin square
	1948(II)	5	Convenient lengths measured subsequently	A followed by C; R, L; I, O	2×2 2×2	8 8	32 and 32	2 Latin squares followed by 2 more Latin squares

For explanation of types of setting see Appendix.

For explanation of C' see Fig. 5 or Table 4.

sual at the beginning of an experiment can be approximated by regression terms; but performance becomes so much less variable, as a subject becomes practised, that it is doubtful whether the assumption made in analysis of variance of the uniformity of residual variance is applicable to the results of an experiment in which a subject shows marked improvement. In the end, however, the results, when factorial designs were adopted, differed from the results of the older classic experiment (which consists of blocks of like responses such as are recommended by Guilford), but learning and adaptation made their appearance in fresh forms.

III. INTER-PERSONAL VARIATION

It will be convenient to discuss the material from two standpoints, with respect first to inter-personal and secondly to intra-personal variation. To take inter-personal variations first, the material falls into two parts, the Scottish material of 1934-8 and the results collected at Oxford in 1946-8.

Table 2. *Scottish data. Inter-personal variation*

Variances in (mm.) ²				
	D.F.	A	B	C
1935-6	13	Systematic		
		4.67 (1st)	35.9 (3rd)	34.5 (2nd)
		Randomized		
1936-7	17	16.9	32.9	41.4
1937-8	17	34.5	43.2	20.7
1938	13	14.9	22.7	25.5

Variance ratios, based on tables (Fisher & Yates (4))

	12		24
	0.05	0.01	0.01
13	2.60	3.96	3.59
17		3.45	3.08

The result which stands out is the much lower inter-personal variance in 1935-6 when judgements of plain lines were made before any judgements with arrowheads or feathers.

The main point which I wish to make about the Scottish material can be clearly seen in the diagrams for the years 1935-8 and in Table 2. In 1935, subjects each made 80 settings of type A, then 80 of type C and finally 80 of type B. They began the experiment, therefore, by making 80 settings with plain lines, as yet undisturbed and undistracted by arrowheads or feathers. In the three subsequent years subjects each made 72 settings in all, 24 each of types A, B and C mingled. The important point is that the inter-personal difference in the mean value for plain lines (settings of type A) is so much less in 1935 than in 1936, 1937 or 1938. Another way of putting the point is that in 1935 the lines in Fig. 1 (each line representing the performance of one subject) tend to diverge from A to C; while in 1936, 1937, and especially 1938, in Figs. 2, 3 and 4 the lines tend to be more nearly parallel. The inference which I draw from these diagrams is that, in some way or another, because of randomization (that is to say, by the mingling of A, B and C settings), subjects seem to have adopted some standpoint about the relation between A, B and C settings and to have made their responses in terms of this standpoint. One might say that the subjects respond to the series of stimuli in some way as a whole, which seems to be another way of saying that a 'gestalt' principle is involved.

There are, beyond this, some rather puzzling features in the data (see Table 3). First, it is not surprising that the mean values for the B settings and the C settings should be

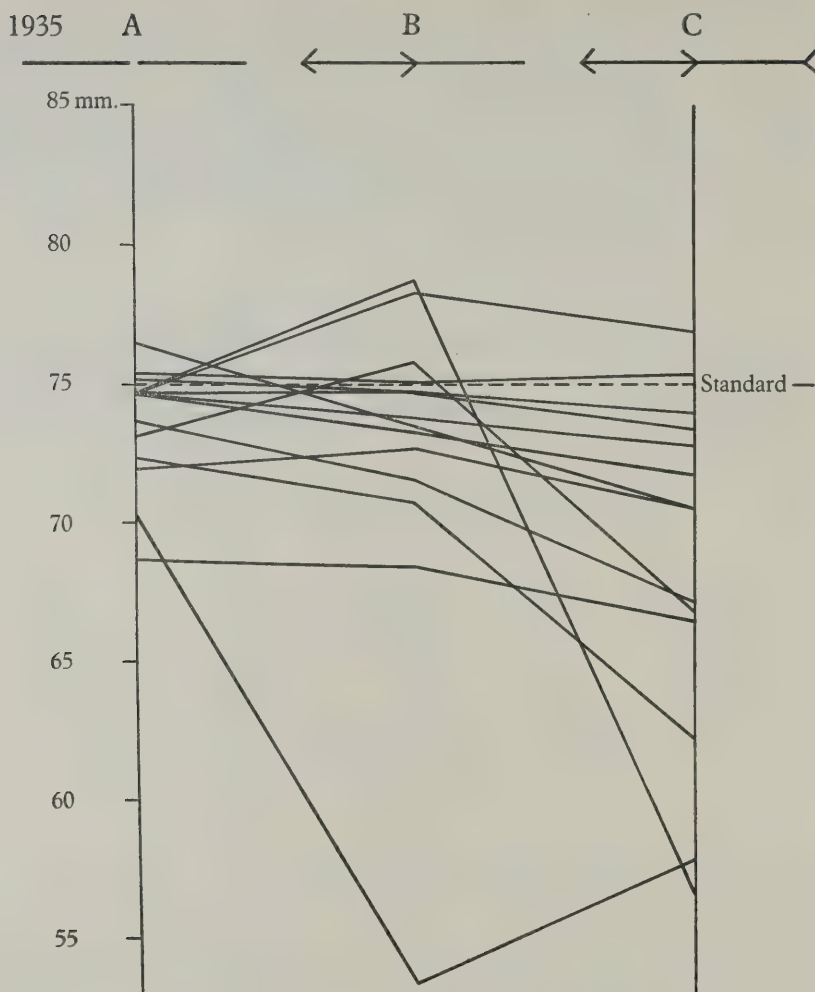


Fig. 1. Divergence after initial stimulus with plain lines.

Table 3. *Scottish data. Means*

No. of readings of each main type	Year	No. of subjects	Deviation from 75 mm. length		
			A	B	C
80	1935-6	14	-1.4	-7.2	-10.9
24	1936-7	17	-2	-5.5	-10.1
24	1937-8	17	-2.2	-4.6	-9.8
16	1938	13	-1.1	-4	-9.4

The residual 'intra-personal' variance seems to be greater when the types of readings are taken systematically i.e. in blocks of like sign, instead of being randomized.

Mean values show little change over the four years except for shift of B across from near C to near A.

less than the standard, but it is at first sight odd that the mean values for the A settings should also be predominantly less than the standard. Secondly it is puzzling at first sight that the mean values for sets of type A should show so little difference from year to year

At the time, this evidence was taken as indicating first, the general comparability between the experimental situation in 1935 and the three subsequent years. While the explanation offered for the defect in the A settings was that the shape of the gadget (see illustration in

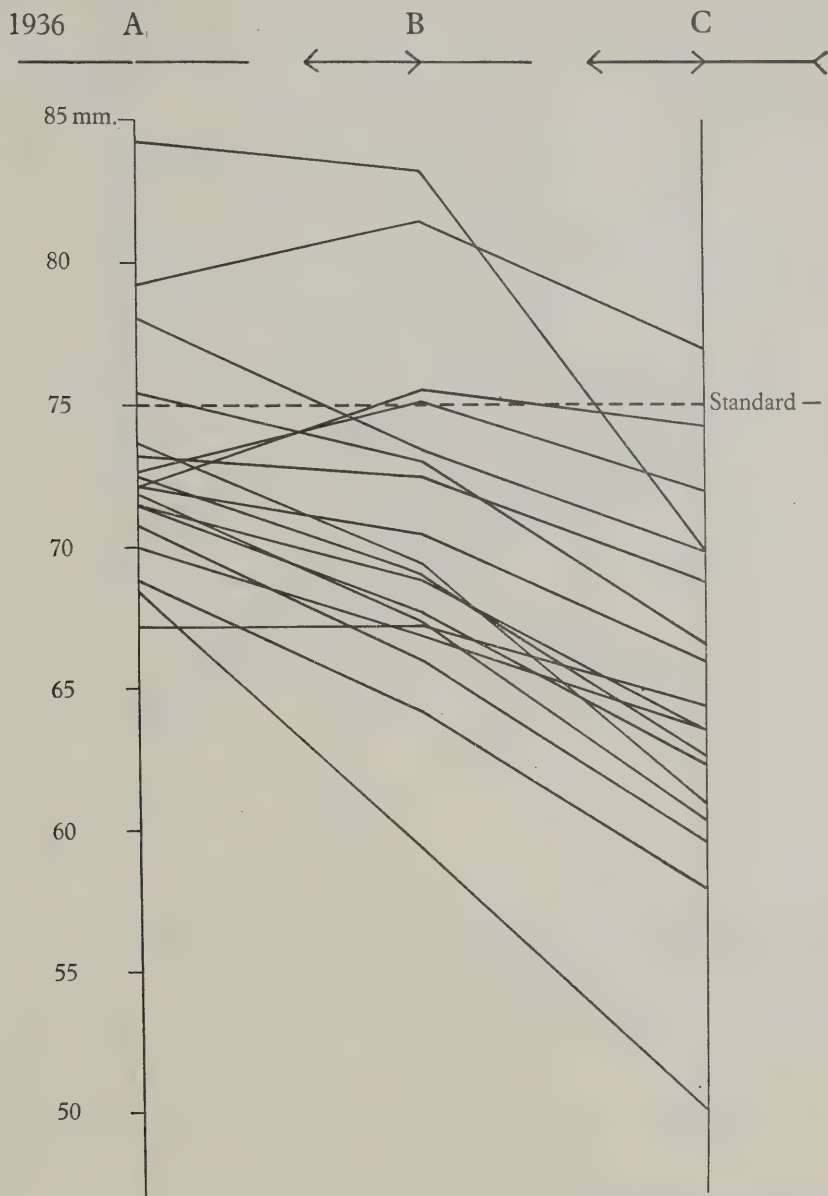


Fig. 2. Divergence after stimulus by A, B and C intermingled.

Appendix) was such that a rather shorter line on the variable slide balanced the fixed standard line on the rather broader sheath, an effect which is quite intelligible as a 'gestalt' phenomenon.

Turning to the Oxford results, it was not possible to run the experiment at Oxford

in exactly the same way, owing to the lack of the standard gadget. I developed instead a system in which observers marked sheets of ruled paper, in advance, with arrowheads or feathers to act as the stimuli. (Sample sheet is shown in Fig. 5.) This procedure saved

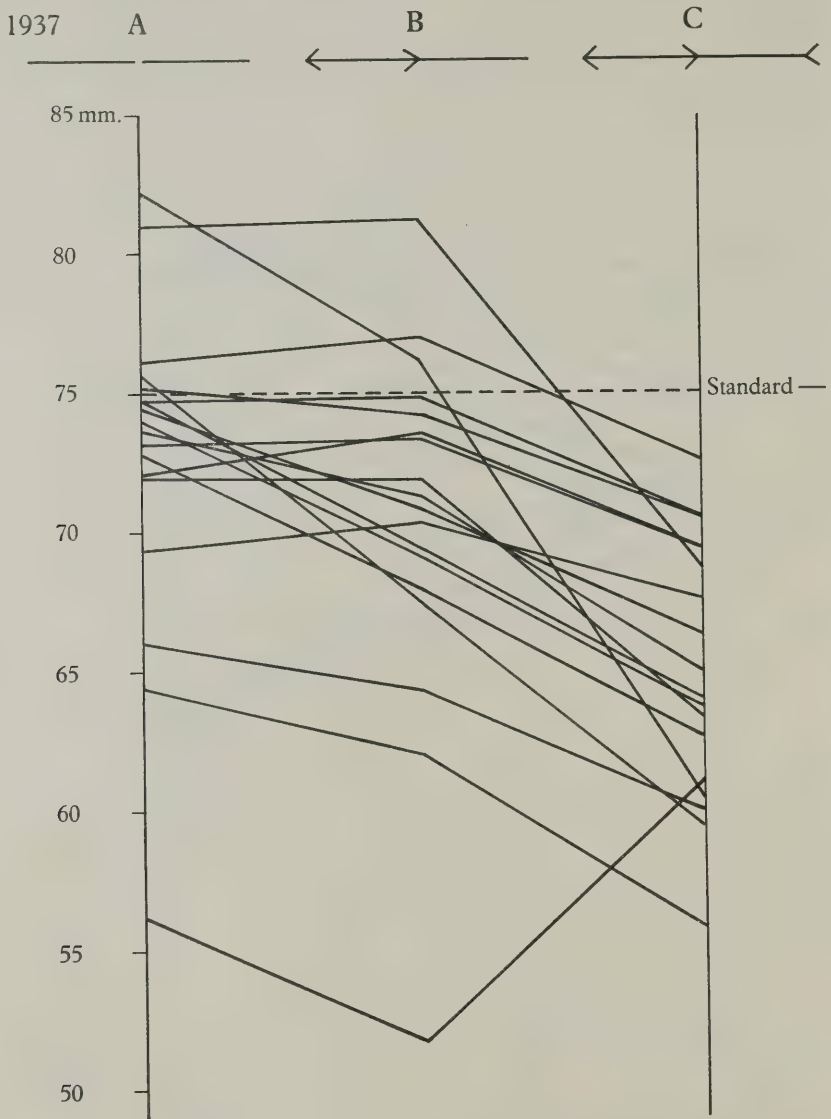


Fig. 3. Divergence after stimulus by A, B and C intermingled.

the expense of apparatus and avoided the fault noted in the Scottish material, namely, the difference of width between the sheath and the slide. On the other hand, it meant that observers either had to spend time drawing lines to exact lengths, or accept the position that the standard lines presented to their subjects would not be of uniform length. On the whole I preferred the latter, but it offended the sense of propriety of some students, and in 1947 the class voted to use prearranged standards, half working with 75 mm. and

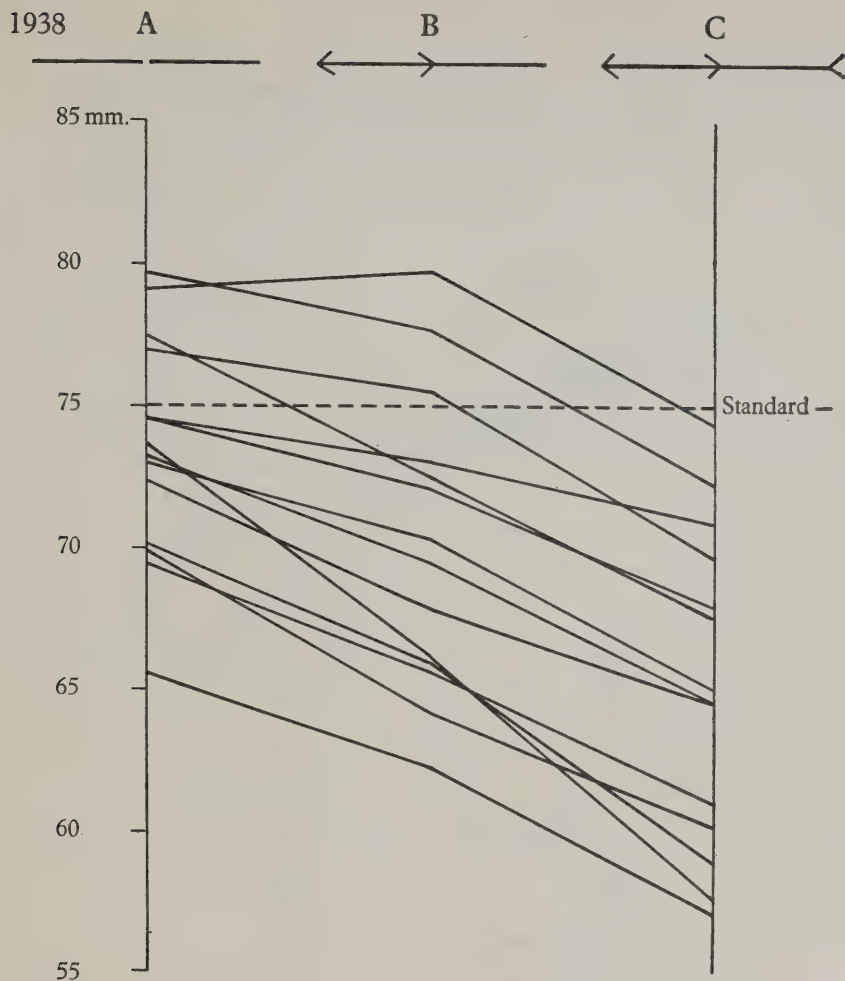


Fig. 4. Divergence after stimulus by A, B and C intermingled.

half with 50 mm. Abrogation of uniformity of standards introduces in turn a fresh problem, i.e. whether to work with deviations as measured or with ratios. For simplicity and to save work in computation, I have used deviations in mm. throughout. There are objections to either procedure; the use of ratios, for instance, involves a deceptive accuracy when both numerator and denominator are of the order of 50 or 75 and are rounded to the nearest mm.

The chief points in the Oxford data to be noted are:

(i) the small inter-personal variance in 1946 for type A (and small mean deviation $+0.2$ mm.);

(ii) some tendency to parallelism in 1946, 1947 and 1948 (I);

(iii) the smaller inter-personal variance of type A in 1948 (II) than in 1948 (I).

These are of interest for the following reasons. The small and positive average personal deviation in 1946 ($+0.2$) was at the time taken as confirming the view expressed that the general tendency for settings of type A to be too short in the Scottish data was due to the type of apparatus. Evidence of other experiments made about the same time, however, suggested the alternative explanation of a central tendency, which would mean here that

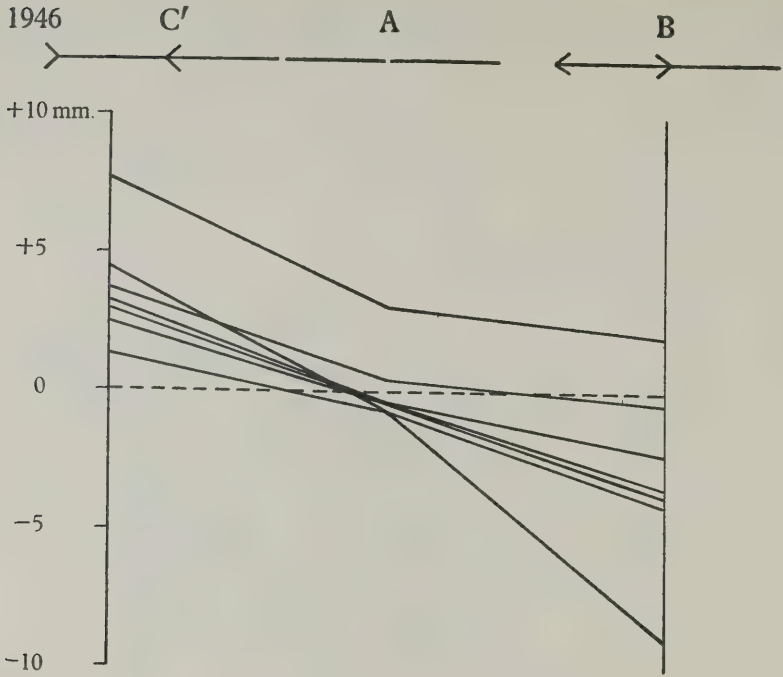


Fig. 6.

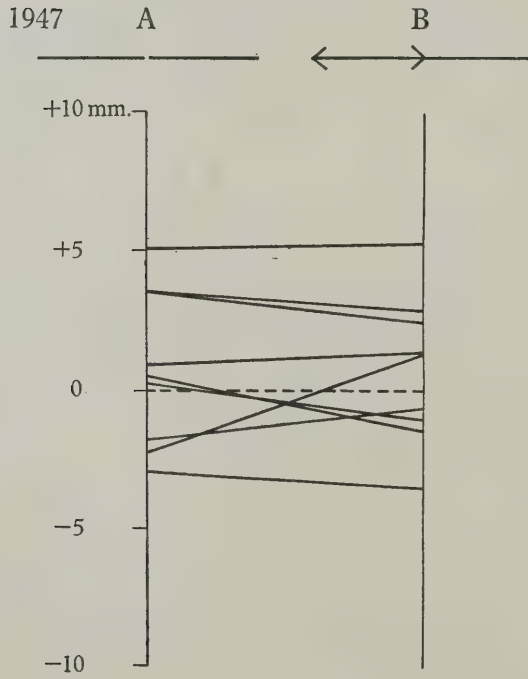


Fig. 7.

Figs. 6 and 7 relate to data collected in Oxford.

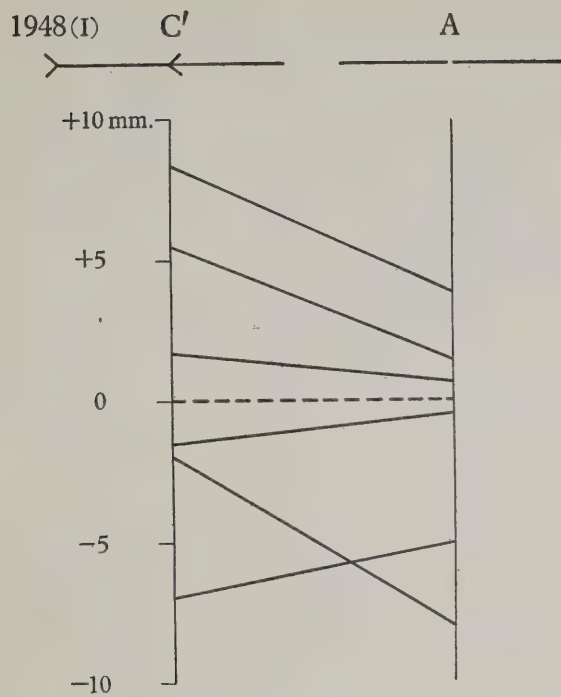


Fig. 8.

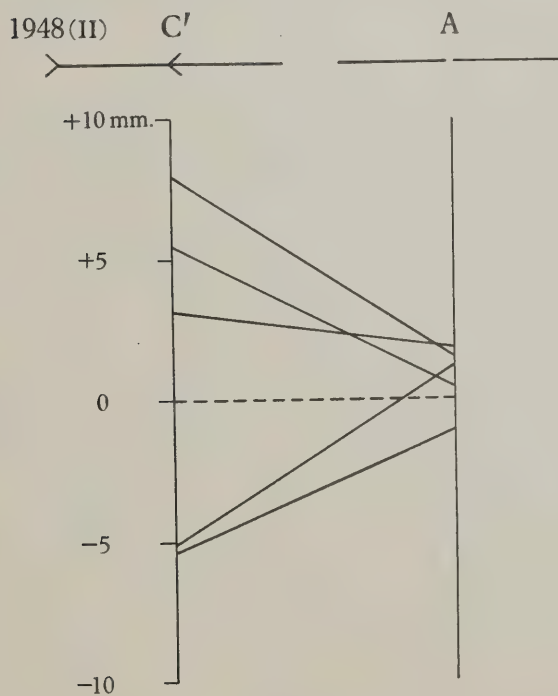


Fig. 9.

Figs. 8 and 9 relate to data collected in Oxford.

Table 4. *Oxford. Inter-personal variation*

Number of subjects	Year	Variances in (mm.) ²		
		C'	A	B
7	1946	1.76	0.84	5.15
9	1947	—	8.22	7.82
6	1948 (I)	20.7	8.7	—
	randomized			
5	1948 (II)	38.9 (2nd)	2.1 (1st)	—
	systematic blocks			

Table 5. *Oxford. Means*

Number of settings of each main type	Year	Excess or defect in mm.		
		C'	A	B
16	1946	+2.5	+0.2	-2
32	1947	—	+0.9	+0.75
32	1948 (I)	+1.6	-0.4	—
32	1948 (II)	+1.1	+0.8	—

the deviations for type A in 1946 were so small because the bias in B settings was in the opposite direction to the bias in the C' settings, whereas in the Scottish data the bias in the B and C settings had been in the same direction. In 1947, therefore, I redesigned the experiment so that bias would operate only in one direction (plain lines A and arrowheads B

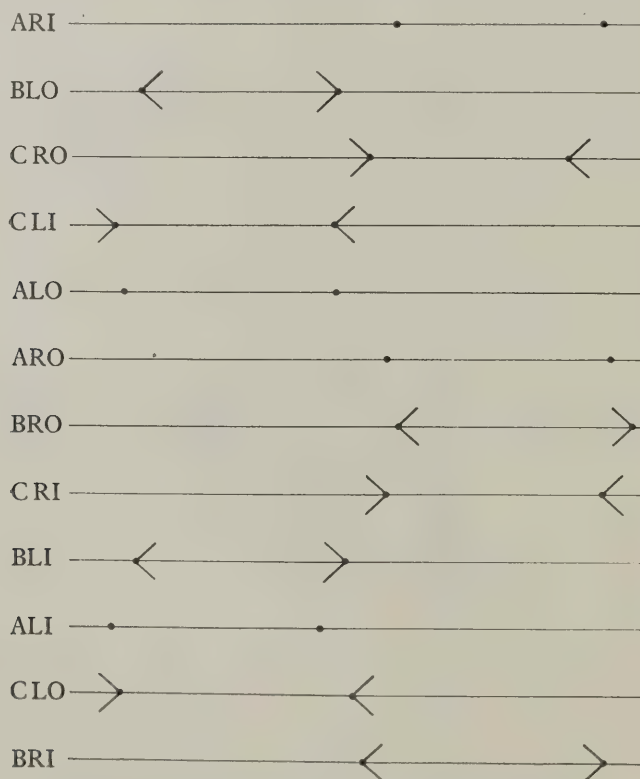


Fig. 5. Sample sheet as laid out for use by Oxford subjects.

only being used). I made the forecast at the time that the defect, which I expected in the B settings, would draw the A settings down below the standard, as had happened in Scotland, using types A, B and C. The failure of this forecast suggested another explanation, that in some subjects the tendency to compensate for the arrowheads was so strong as to produce a mean excess instead of defect and to set up a 'gestalt' carrying the A settings with them. If this explanation is acceptable, the 1948 (I) results become quite intelligible. This time I used two types of setting: plain lines A and divergent feathers C, expecting a tendency to make the variable line too long. But with the evidence available, I now looked for parallel lines on the diagram, with no knowledge as to whether they would be predominantly above or below the standard. The actual result now seems to depend on the state of naïvety of the subjects and the extent to which they over-compensate. Furthermore, the curious point noted in the Scottish data, that the mean values for A settings were approximately the same when mingled with B and C settings as they were in isolation, becomes intelligible. If they had met the illusion before, subjects had been deliberately instructed not to attempt to maintain a naive approach, and I suggest therefore that, where a subject deliberately compensated for the arrowheads, this tendency carried over to the A settings as well—a situation which may be conveniently expressed by saying that the subjects were responding to a 'gestalt' which developed, rather than to a series of isolated stimuli.

It remained to see if, under the Oxford conditions, the inter-personal differences would be reduced if settings of type A were taken in isolation as they had been in Scotland in 1935. Confirmation is indicated by the results of 1948 (II) as shown in the diagram, but unfortunately the number of cases is very small.

IV. INTRA-PERSONAL VARIATION

This can be broken down into (i) response to changes in the factorial combinations, (ii) progressive changes (the design was in each case intended to allow short-term changes to be distinguished from longer term changes by balancing successive blocks of stimuli), and (iii) residual or 'error' variance. Full analysis of intra-personal variance would overweight this paper, and the matter must be dealt with summarily.

(i) *Response to main factors*

Marked intra-personal variations in response to the main factors are noted, but there is little that calls for detailed comment. Some people were much more susceptible to the illusion than others, but then there were undoubtedly differences in the familiarity of subjects with the 'Muller-Lyer illusion'. Space effects, i.e. differences between working with the standard on the right or on the left, were rather surprisingly large in certain cases, but the direction of the effect seems to be distributed quite indiscriminately, about as many subjects making the variable longer when it was on the right as did when it was on the left. There seems to be a systematic difference between settings made by increasing the length of the variable line and that made by decreasing it; the predominating tendency has been to stop short, or too soon. As a result, the averages for settings made by decreasing the line are in the great majority of cases greater than those made with increasing lines. This result is hardly surprising, and the 'movement error' has often been noted before in psychometric experiments.

(ii) *Progressive changes*

In most cases the long-term changes, that is to say, differences between successive blocks of readings, are more important than short-term changes and there is rather a large number of cases in which the variance attributable to short-term changes is less than the 'error' variance—an observation which suggests some sort of negative correlation within blocks. Attempts to fit linear regression lines to short-term changes yielded little of interest.

(iii) *Residual or 'error' variance*

The results (see Table 6) are puzzling and the picture is obscure. Two possibilities were considered, the first that by randomizing the stimuli and then analysing the variance, one would be left with a smaller residual variance. The second was that, because of randomization, the subjects would be prevented from assuming a stable state and that the residual variance would therefore be increased (8).

Table 6. *Residual variance in mm.²*

Year	No. of subjects	Standard	Average variance	Arrangement of factors
1934-5	22	75 mm.	9.36	Systematic
1935-6	14	75 mm.	6.74	Systematic
1936-7	18	75 mm.	11.92	Randomized
1937-8	5*	75 mm.	8.65	Randomized
1938	14	75 mm.	9.4	Randomized
1946	6	Variable	8.14	Randomized
		c. 4-7 mm.		
1947	4	{ 75 mm.	8.05	Randomized
		{ 50 mm.	5.93	Randomized
1948 (I)	5	Variable	7.07	Randomized
		c. 4-7 mm.		
1948 (II)	5	Variable	7.7	A only, C only, R.L.I.O.
		c. 4-7 mm.	10.38	randomized in each

If anything, the evidence points to an increase in the residual variance, but the difference is not uniformly in this direction. Comment on this is made in the next section on qualitative observations.

V. QUALITATIVE OBSERVATIONS

The data were collected in the form of reports by students, who acted as observers and subjects. Most of the reports have been returned to the writers; introspective and observational material is therefore, in the main, no longer available, and it is in any case doubtful whether either was at all fully treated in the reports themselves.

In the course of supervising the experimental work, I have been continually surprised at the difficulty of ensuring that a procedure should be carried out uniformly by a group of student observers. Two aspects of this difficulty may be mentioned: first, that glaring departures from the instructions were committed and tolerated by students, and secondly, the difficulty of showing that such departures have in fact affected the results. Two recent examples spring to my mind: X, who refused to allow his observer to move the card along the variable line, as laid down in the manual, but insisted on doing it himself, and Y and Z, who sat opposite one another, each with a sheet of stimuli, acting as observer or subject for alternate stimuli.

It is always of interest to see to what extent subjects' performances fit in with previous ideas about them. Generally speaking, the residual variance has reflected pretty well the impression of subjects derived from their other work, as being careful or carefree. Response

to the main effects and interactions has been much less obviously connected with such impressions of personality and it has only been in exceptional cases that a result squared with, or made sense in relation to, impressions of the subject which I had formed.

Perhaps the most difficult point to accommodate in a statistical model is the feeling that subjects are apt to develop a trend, i.e. will give a series of responses which become progressively longer till suddenly S recognizes what he is doing and the next response jumps in the other direction. Low-order polynomials are inadequate for this type of trend. I am inclined to take the view that autocorrelation is inadequate too; where I have worked out correlations between successive residuals, they have been extremely low. There is, moreover, a moral certainty that no subject maintains throughout an experiment the same level of dependence between responses and their predecessors. The situation is, in fact, that we know almost for certain that each response is influenced by its predecessors, that we can seldom if ever assess the extent of this influence, and that an overall statistical assessment will be quite powerless to represent it. It is this situation, where responses are psychologically related but statistically independent, which is so elusive. I seriously question the value of attempting to assess the exact relation between a response and its predecessors, because from what I have seen of psychological experiment, such a relation seldom has any permanent form.

Descriptively, the process seems to be something like this: in the early stages the subject attempts to judge each stimulus on its merits—whether this is achieved is another matter, but, by and large, one can say this is his intention. Quite soon, however, the subject seems to find that the present evokes comparisons with the past, not necessarily the immediately preceding stimulus and response, but past stimuli and responses previously made. This experience is accompanied by, or gives rise to, attempts to compensate for features of the present situation. Another recognizable stage, which usually occurs early in proceedings, is an attempt to devise a system, though I do not think I have ever found a subject who persisted with a system throughout. There are many misleading systems, e.g. any system based on duration of movement is apt to involve the 'stimulus error' in some form. The situations in which I have observed this failure of a systematic approach have all been conducted without the subject knowing his results.

These stages merge into yet another, in which the subject finds less and less reference to an explicit past and the responses are in some sense known immediately. I find the description of this stage is very difficult, even in my own case. It may only be a rationalization to say that I am relying on past experience or that I have learnt, but it is hard to escape the impression that the responses are made in terms of some aggregate of, or central tendency of, past experience. In other words, an internal pattern seems to have developed among stimuli of the several kinds employed. It is this situation which, I suggest, may be described as the development or emergence of a 'gestalt' and which is illustrated most clearly in the diagram for 1938. The terms 'development' and 'emergence' are used advisedly in spite of the usual view that 'gestalt' qualities are immediate; they seem appropriate because there is fairly strong evidence of some sort of patterning by the end of an experiment, while it is very difficult to see how a psychological pattern can, in any important sense, exist before the experiment begins. To take an example from a different experiment, in which the subject had to attempt to match a number of small scattered areas with one continuous area, the change was from 'this stimulus is larger than the last' or 'this stimulus was the one I saw second in the last series' to 'I know this stimulus matches

area number 3'. Objectively, the level of response may change even now, but subjectively development is very slow once this point is reached.

This presents a very condensed and simplified version of what seems to take place. I will not maintain that the four steps occur recognizably in all subjects nor that they last as long for different subjects. Personally, I think that the autoregressive model is inadequate for any experiment long enough to cover the three stages I described, because to some extent the differences seem to be qualitative. To bring out the nature of these difficulties more clearly I have made a speculative attempt to devise a statistical model. In view of what subjects say, and what I have experienced myself, any regression model based on the effects of the last response or the last two or even the last few responses, is almost certainly inadequate. I doubt if the following suggestions would be adequate or if they offer a practicable basis for computation. My point is that any model would have to be *at least* as complicated in order to take account of observed phenomena. Perhaps it could be done if we envisaged a model in which the number of terms in the regression equation increases each time a response is made. But I think the regression coefficients on the nearest terms would decrease at the same time; in other words, the subject would tend to respond in terms of a progressively more stable aggregate or average of past responses, say

$$R_{n+1} = n^a, R_n + n^{a_2} R_{n-1}, \dots,$$

$$R_{n+2} = {}_{n+1}a_1, R_{n+1} + {}_{n+1}a_2 R_n,$$

where R_l is the l th response.

It is important, however, to recognize that this is not the whole story. Such an equation could I think represent the progress from a guess to a judgement in terms of some pool or aggregate of past experience, but not the qualitative changes implied by the second and third stages which I have described above. From the subjects' reports and my own observations we must I think consider two more influences. First, the subject's expectations about the stimuli; for a subject brings with him to an experiment his own personal expectations of the sort of response which he is liable to make or it is reasonable to make. Thus, he does not expect, in an experiment in which he is told to say whether the weight in his right hand is heavier or lighter than the one in his left, to be saying 'heavier' all the time. Subjects are inclined to become restive if they find themselves responding in a way perceptibly at variance with their expectations. They tend to slip in a few responses in the other direction if there is any excuse, or perhaps even to restore their self-respect.

This aspect of response has of course considerable importance for the principle of randomization. Without going into detail, I throw out the suggestion that the set of orders of a series of digits or other groups of variates which are acceptably random to a human subject are not necessarily all those which satisfy the usual statistical tests. Whether some will satisfy a human being which would fail at the statistical proof I am not sure, but I don't deny the possibility. A term $P_s(R_n \cdot R_{n-1} \cdot R_{n-2} \cdot R_{n-3} \dots)$ is required, where P_s means the 'subjective' probability, that is, the probability or credibility for the subject that the stimuli should have occurred, as he supposes them to have occurred. R_j is the j th response, and the 'weight' of a response varies with its recency.

The other factor I have in mind is the subject's general level of threshold for variations in response. It is not difficult to find examples of cases where part of a sequence of responses suggests a trend followed by a sharp change. Subjectively, the subject suddenly says to himself 'that last or those last responses are too long or too short; the next one must be

different'. Roughly speaking there is a range on either side of the response corresponding to the intention of the subject within which variations are not normally perceptible. But if by chance there is a run of deviations in one direction, and particularly if they grow progressively in one direction till they fall outside the threshold, the subject is liable to wake up suddenly to what has been going on and to make some relatively violent correction. This type of control is notorious in current literature as 'feed back' and a suitable term could be written into the equations. It is again of the form $P_s(R_n \cdot R_{n-1} \cdot R_{n-2} \dots)$, the earlier terms carrying less weight. But I suggest that, here, only the more recent terms carry appreciable weights.

In the present and many other psychological experiments, responses are made and recorded; they are not open to review and the subject has no knowledge of results. These restrictions probably affect the subject's outlook. A statistical model of a situation where the subject was free to revise responses, and had knowledge of success, might well be still more complex.

If the subjective effects of the arrowheads and feathers in this experiment remained constant throughout, presumably the equations which I have discussed would refer solely to the residual terms derived from the original responses after allowing for the factorial combination for each. Part of my thesis, however, which the diagrams are intended to support, is that the main effects are not stable. There seems to me evidence that in a sequence of responses the results may be described by saying either that they interact, or by saying that there must be some sort of regression terms for the main effects too. Thus I do not conceive of a subject being susceptible to the same degree to the effect of an arrowhead at the end of an experiment as he was at the beginning. Subjects deliberately attempt to allow for such influences and in practice seldom make a just allowance. But in some sense they settle down to a relatively stable allowance, and it is this balance which seems to develop which is as interesting as anything in the material, and which seems to warrant the use of the term 'gestalt'.

Psychological experiments such as those covered in this paper, which attempt to determine measures as statistical averages, involve a dilemma over and above the difficulties I have discussed, which is hard to escape. Either we present subjects with a situation and attempt to allow for the rate at which they become adapted, or we train them in responding to the situation in question till adaptive changes can be treated as negligible. The former method means statistically much more complicated procedures, while the latter means that the data we obtain are always different from the naïve responses of the man in the street.

VI. DISCUSSION

The defects of such a casual method of experiment (the collection of results of class experiments) are glaring. Many comparisons of interest cannot be made. The lack of uniformity of procedure is responsible for an unknown amount of variation. At the same time I suggest that (overall) a sufficiently strong picture has been built up to claim attention and to deserve refutation or confirmation.

If the main theme has validity, it is difficult to see how to design experiments which can produce assessments of the main effects in isolation. It would amount to this, that we cannot allot a permanent effect, say 10% to arrowheads, or a particular length at a particular angle, but can only say that in certain circumstances the effect was 10%. If the

judgement or setting by the subject is led up to differently or if his expectations are different, then the quantitative effect may be different too.

It is nothing new to psychologists to hear that a response is made to a 'behavioural' rather than to a 'geographical' situation, or in other words, that a 'gestalt' develops. What may have a certain novelty is that for a subject faced with a factorial design of randomized stimuli the behavioural situation may become patterned or develop a 'gestalt'. The novelty should not prevent one, moreover, from seeing that the behavioural situation which develops with one randomization of a factorial design may quite easily differ from that which develops with another.

The question of significance tests well merits further study. Any dependence of residual errors on the design of the experiment seems to me to strike at the root of significance tests and in particular at any tendency to regard reduction in size of residual error by itself as a sign of efficiency of experiment.

VII. SUMMARY

Central among the assumptions underlying analysis of variance are, first, that the effects involved are additive and, secondly, that observations are independent of one another. The author holds that both assumptions are suspect in psychological situations of the kind where a human subject is required to make a number of responses. The evidence presented in support of this view is drawn from a particular experimental setting, but in the author's experience there is reason for regarding the dangers as general and active over a very wide field.

A class experiment, designed on the classical scheme of altering one variable at a time, had been undertaken with 'Muller-Lyer' apparatus to obtain average values for effects of various arrangements of arrowheads. Initial results were promising enough for further series of class experiments to be undertaken with the object of showing the efficacy of factorial designs and the usefulness of analysis of variance with psychological material. The possibility was also envisaged that randomization of the order in which various factorial combinations were presented would help to prevent or delay learning or adaptation. The results obtained point to certain unexpected and fundamental difficulties in interpretation.

It is argued that the effect of a factorial design, according to which a subject makes a series of responses under a variety of conditions, is to set up a frame of reference peculiar to the design. Thus the response to a particular combination of controlled conditions depends to some extent on the factorial design in which it occurs. Another way of putting this is to say that when a factorial design is used a 'gestalt' emerges to which the subject responds.

Qualitative evidence strongly suggests that successive responses (at any rate when made, as here, at short intervals) are dependent in varying degrees on the stimulus-response situations which have preceded them. An attempt is made to describe the more important forms of dependence. It is suggested that linear or low-order regression terms would not be adequate.

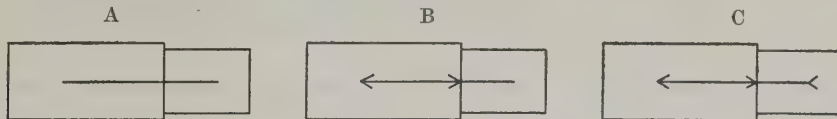
Evidence as to whether estimates of residual, or error, variance are reduced by the use of factorial designs is inconclusive.

The conclusion is that, by the nature of the case, at least two fundamental assumptions of analysis of variance are not satisfied, and factorial designs no longer carry their peculiar merits. Interpretation of results by the usual methods and standards is to be regarded as dangerous in psychological situations where a subject makes a series of responses, and in the interpretation of results obtained in such a way.

APPENDIX

(a) Extract from the manual used for the class experiments 1936 and 1937.

(i) The apparatus will be used in three forms:



(ii) The apparatus may be presented to S with the slide to the right (R), or the left (L) of the sheath.

(iii) The apparatus may be presented to S with the slide pushed in requiring to be pulled out (I), or already out and requiring to be pushed in (O).

Thus there are twelve possible ways in which the apparatus may be presented to the subject, namely:

ARI, ARO, ALI, ALO; BRI, BRO, BLI, BLO; CRI, CRO, CLI, CLO.

Investigator and S make four trial settings, say, ARO, BLI, BRI, CLO, to familiarize themselves with the procedure. The results of these settings should be recorded on the side of the record sheet.

Then proceed to the main experiment. This consists of six series of twelve settings each as shown in the following scheme, *which is to be followed carefully*.

Order of presentation	Series					
	I	II	III	IV	V	VI
1	CRI	BRI	ALO	CRO	CLO	BRO
2	CLI	ARO	BLO	ARI	BLI	ALI
3	ALI	ARI	CRI	ARO	ALO	CRO
4	BRI	BLI	CLO	BLO	BRO	CLI
5	ALO	CLI	BRI	BLI	CRI	CLO
6	BLO	ALI	ARI	BRO	CRO	ARO
7	ARI	ALO	CRO	BRI	BLO	CRI
8	BRO	CLO	ALI	CLI	ARO	BLI
9	BLI	CRI	ARO	ALO	ARI	BLO
10	CLO	CRO	BRO	ALI	CLI	BRI
11	CRO	BLO	CLI	CLO	ALI	ARI
12	ARO	BRO	BLI	CRI	BRI	ALO

The apparatus should be laid on a neutral background so that S looks directly at the line dividing the slide from the sheath.

Investigator gives the instructions, 'Hold the sheath firmly and set the slide so that the two lines are equal in length'.

REFERENCES

- (1) BRUNSWIK, E. (1949). Systematic and representative design of psychological experiments. *Berkeley Symposium*. Ed. J. Neyman. University of California Press.
- (2) COCHRAN, W. G. (1939). Long-term agricultural experiments. *J. R. Statist. Soc.* (Suppl.), vi, 104.
- (3) EDWARDS, ALLEN L. (1950). *Experimental Design in Psychological Research*. New York: Rinehart and Co. Inc.
- (4) FISHER, R. A. & YATES, F. (1948). *Statistical Tables for Biological, Agricultural and Medical Research*. Edinburgh: Oliver and Boyd.
- (5) GUILFORD, J. P. (1936). *Psychometric Methods*. New York and London: McGraw Hill Book Co. Inc.
- (6) KATZ, D. (1949). Gestalt laws of mental work. *Brit. J. Psychol.* xxxix, 175.
- (7) KOFFKA, K. (1935). *Principles of Gestalt Psychology*, p. 682. London: Kegan Paul.
- (8) MULLER, G. A. & GARNER, W. R. (1944). Effect of random presentation on the psychometric function. *Amer. J. Psychol.* lvii, 451.
- (9) SUMNER, W. L. (1948). *Statistics in School*. Oxford: Basil Blackwell.
- (10) WALSH, J. E. (1947). *Ann. Math. Statist.* xviii, 88.

(Manuscript received 6 May 1950)

ATTENTION TO CHANGE

By D. E. BERLYNE, *University of St Andrews*

I. *Introduction* (pp. 269-270). II. *Experiments* (pp. 270-273). III. *Discussion* (pp. 273-277). IV. *Summary* (p. 277). V. *References* (p. 278).

I. INTRODUCTION

The power of changing or moving objects to attract attention has long been noted in psychological primers and exploited in practical life. Among a multitude of instances, one can cite the use made of this power by designers of advertising signs, and the recourse of many animals, at times when conspicuousness means danger, to lying low and remaining inert. But sensitivity to change and novelty, whose biological necessity may incline us to take it for granted, is not at all easy to account for completely in psychological or neurological terms.

A previous paper by the writer⁽²⁾ suggested a method of studying attention through observation of overt behaviour. This method was used in a series of experiments on the influence of *stimulus intensity* on attention, and had the advantage of showing this influence to be deducible from the principles of behaviour theory. The present paper reports an attempt to bring the same experimental and theoretical techniques to bear on the problem of *stimulus change*.

Such words as 'change' and 'novelty' are often employed loosely to cover a wide range of factors which need to be carefully distinguished, e.g.:

(1) A stimulus may be 'new' or 'unusual' in relation to *an individual's total life-history*, and such a stimulus is likely to attract attention. In a previous study⁽³⁾ we have given reasons for believing that 'completely new' stimuli of this sort arouse a drive-stimulus-producing response. The intention was to relate curiosity to the conception of fear or anxiety recently elaborated by Miller and Mowrer^(17, 18, 19, 20), except that these writers are thinking of secondary or acquired motivation, whereas this type of curiosity appears to be an innate tendency. The exploratory activity with which animals allay their curiosity can explain without much difficulty the effect on attention⁽⁴⁾.

(2) A stimulus may be 'new' in relation to *what has immediately preceded it*.

This is the state of affairs to which the word 'change' is properly applied, and includes 'suddenness' and 'movement', mentioned as determinants of attention by Titchener⁽²³⁾, p. 209). To explain these, he contents himself with the observation that 'given the course of development as we know it, the organism must have attended to movement, etc., in its surroundings or have paid the penalty of inattention with its life. The moving, the new and the sudden are all possible—even probable—sources of danger.' This accounts for the biological utility of the effect in question, but evades the problem of the mechanism by which it is brought about. It would seem advisable to distinguish special cases with somewhat different psychological implications:

(a) The influence of *perceived movement* is not surprising, in view of the low threshold for movement as a perception *sui generis*.

(b) Similarly, sudden changes affecting the activity of particular afferent fibres may be

expected to act in much the same way as increases of intensity, since the responsiveness of nerves to rates of change as a result of adaptation is well known⁽¹⁾.

(c) *Surprise*, or conflict between a perception and an expectation, is a factor whose capacity to draw attention is demonstrated in experiments by Wilcocks⁽²⁴⁾ and by Desai⁽⁶⁾. Both these investigators used an increase in reaction-time as a sign that attention is being appropriated by the unexpected features of a stimulus, and Desai used in addition disturbance of a continuous motor activity. In some conditions, however, Wilcocks found that an unexpected change made something *less* likely to be noticed. A theory to account neurologically for the emotional effects of unconfirmed expectancy is offered by Hebb⁽⁹⁾, pp. 149–50), and may allow us to see in the relation between surprise and attention merely a special case of that obtaining between motivation in general and attention⁽⁴⁾. (This is assuming that emotions function as motivational forces, as Mowrer reasons⁽²⁰⁾, and using the word 'attention' in a different sense from Hebb's 'perceptual set' (⁽⁹⁾, p. 102).)

(d) But an altogether different and less simple problem is presented when the change from old to new is not experienced as a momentary event, because the stimuli are intermittent, and when the effect on attention persists for some time after the change has occurred. Such cases form the subject-matter of this paper. We shall first describe experiments demonstrating that attention is affected by this kind of change, and then we shall consider the significance of the phenomenon for behaviour theory.

II. EXPERIMENTS¹

Apparatus

The apparatus and method used were in essentials the same as those described in detail elsewhere⁽²⁾. It will, it is hoped, suffice to recapitulate briefly that the apparatus presented the subject with one or more at a time of four visual stimuli, whose colour and shape could be varied. Each stimulus had its corresponding Morse key, which was to be pressed in response to it; when a combination of more than one stimulus appeared, the subject was to respond to one only, the choice being left free. Response to a particular stimulus was regarded as an objective manifestation of attention to it.

Subjects were students, naval ratings and R.A.F. air crew.

Procedure

(1) *Experiments 1–3*

The general design was as follows (for details see Table 1):

An experimental group was first (Phase I) given a series of forty-eight stimulus-patterns (ones, twos and threes in random order) selected from four homogeneous stimuli. In a second and a third series (Phase II and Phase III) one of the four was replaced by a different stimulus. The number of responses made to this new and odd stimulus when it was accompanied by one or two of the others, was compared with the number to be expected by chance. Only the trials involving two or more heterogeneous stimuli were counted as relevant to the problem. The others were intended, together with rotation and randomization, to minimize the effects of expectancy, positional preferences, etc.

A control group underwent the same procedure with the exception of the preliminary phase involving four homogeneous stimuli. Phases I and II used, therefore, patterns

¹ These experiments were carried out early in 1948 at the Cambridge Psychological Laboratory with the help of a Rockefeller Research Grant.

selected from three stimuli of one class and one stimulus of another, exactly as in Phases II and III, experimental group. The control group verified that any above-chance tendency to respond to the changed stimulus was due to its having changed and not to the fact that it was in a minority or some other variable (e.g. differences in brightness, area or colour). The three experiments tested the effects of different kinds of change as follows:

Table 1. *Experimental design*

Experiment	Group and no. of subjects	Phase	Stimuli (changed or odd stimulus in italics)
1 (change in shape)	Experimental (16)	I	4 white circles
		II	3 white circles 1 <i>white square</i>
		III	3 white circles 1 <i>white square</i>
		I	3 white circles 1 <i>white square</i>
		II	3 white circles 1 <i>white square</i>
		III	3 white circles 1 <i>white square</i>
	Control A (16)	I	3 white circles 1 <i>white square</i>
		II	3 white circles 1 <i>white square</i>
		III	3 white circles 1 <i>white square</i>
	Control B (16)	I	4 white circles
		II	1 white circle 3 <i>white squares</i>
		III	1 white circle 3 <i>white squares</i>
2 (change in colour)	Experimental (9)	I	4 white circles
		II	3 white circles 1 <i>red circle</i>
		III	3 white circles 1 <i>red circle</i>
	Control (9)	I	3 white circles 1 <i>red circle</i>
		II	3 white circles 1 <i>red circle</i>
		III	3 white circles 1 <i>red circle</i>
3 (change in colour)	Experimental (8)	I	4 red circles
		II	3 red circles 1 <i>white circle</i>
		III	3 red circles 1 <i>white circle</i>
	Control (8)	I	3 red circles 1 <i>white circle</i>
		II	3 red circles 1 <i>white circle</i>
		III	3 red circles 1 <i>white circle</i>
4 (periodic and progressive change)	Experimental (12)	I	4 red circles
		II	3 red circles 1 <i>white circle</i>
		III	3 red circles 1 <i>white square</i>
		IV	3 red circles 1 <i>green square</i>
		V	3 red circles 1 <i>green circle</i>

Experiment 1 dealt with change in *shape*. The changed stimulus was a *white square* (approx. $\frac{1}{2} \times \frac{1}{2}$ in.) among *white circles* (approx. $\frac{1}{2}$ in. in diameter). A second control group (control group B) repeated the procedure of the experimental group, but with *three* changed stimuli in Phases II and III. This was to ascertain whether the changed stimuli must be few in number for the effect to occur.

Experiment 2 was intended to show whether the preference for the changed stimulus appeared after change in *colour*. The changed stimulus was a *red circle* among *white circles*.

Experiment 3 was concerned also with change in colour, but tested still further the hypothesis that change was responsible for the preference by reversing the roles of red and white. The changed stimulus was a *white circle among red circles*.

(2) *Experiment 4*

This experiment showed the effect of a stimulus that changes *periodically and progressively*. After a preliminary series of forty-eight responses (Phase I) with four *red circles*, there followed four series of twenty-four (Phases II–V), in which one of the four was replaced by a *white circle*, a *white square*, a *green square* and a *green circle* in turn. There was no control group. Instead, the number of responses to the changing stimulus was

Table 2. *Results*

(NS = not significant)

Experiments 1-3							
Experiment	Group	Phase	Mean no. of	Mean no.	<i>t</i>	<i>df</i>	<i>p</i>
			responses to changed or odd stimulus (<i>O</i>)	expected on null-hypothesis (<i>E</i>)			
1	Experimental	II and III	13.94	12	2.45	15	0.027
		II	7.5	6	2.4	15	0.03
		III	6.44	6	NS		
	Control A	I and II	12.82	12	NS		
	Control B	II and III	12.31	12	NS		
2	Experimental	II and III	22.22	16	2.35	8	0.047
		II	11.78	8	NS		
		III	10.44	8	NS		
	Control	I and II	15.11	16	NS		
3	Experimental	II and III	24.125	16	5.04	7	0.002
		II	12.125	8	2.98	7	0.022
		III	12	8	3.75	7	0.007
	Control	I and II	12.125	16	2.96	7	0.022
(O less than E)							

(*O* less than *E*)

Differences between means for experimental group and control group

Experiment 1: NS

Experiment 2: NS

Experiment 3: $t = 5.98$, $df = 14$, $p < 0.001$.

Experiment 4

We take the mean number of responses to the changing stimulus in Phases II–V and halve it (i.e. we take the mean per 48 responses). We then compare this with the mean number of responses to the *corresponding* red circle (i.e. the stimulus in the same position as the changing stimulus) in Phase I.

The mean increase = 3.9. $t = 3.84$, $df = 11$, $p = 0.0026$.

Five subjects made the maximum possible number of responses (i.e. 18) to the changing stimulus in Phase II and four of them continued to do so for the rest of the experiment.

compared with the number made to the corresponding stimulus (i.e. the one in the corresponding position) in Phase I.

Results are shown in Table 2. They may be summarized as follows:

(1) In all experiments there was a significant tendency to respond to the changed stimulus in the experimental group rather than to these remaining unchanged.

(2) This tendency was absent in every control group (including control group B Experiment 1), so that it must be due to the change.

(3) The tendency was particularly marked when the stimulus changed periodically and progressively (Experiment 4).

(4) The tendency appeared to decrease with time. In Experiments 1–3 there were fewer responses to the changed stimulus in Phase III, experimental group, than in

Phase II. (In Experiments 1 and 2 the number is significant in Phase II but not in Phase III, while in Experiment 3, it is significant in both phases.)

III. DISCUSSION

The experiments would thus appear to have confirmed the hypothesis that a stimulus which has recently changed is more likely to be responded to than others which have remained unchanged for some time. There is no difficulty in thinking of countless everyday occurrences of which our experimental situations form simple prototypes: we are prone to notice immediately a change in the furniture of a well-known room or a new feature in a familiar landscape. In all these cases the changed element need not be completely new to us, but may have been experienced time and time again in different contexts; it is an innovation only with respect to what preceded it.

However, it is widely agreed that the function of science is not merely to discover new facts or to verify familiar ones, but also to integrate and explain them, i.e. to show how they can be deduced or predicted from basic principles (universal implicative propositions) from which many other empirical data have also proved derivable. Our first step towards this is to consider one by one certain factors which may be ruled out without more ado:

- (1) *A preference for a particular colour or shape* is excluded by the use of control groups, demonstrating that the preliminary phase with homogeneous stimuli is necessary for the effect to occur. It is further checked by the mutual reversal of red and white in Experiments 1 and 3.
- (2) *Positional preference*, such as the tendency to attend to the middle two stimuli, is excluded by rotating the odd stimulus among the four positions for different subjects.
- (3) *A tendency to respond to an odd stimulus* (cf. Wilcocks's experiments on heterogeneity⁽²⁴⁾) is excluded by the use of control groups, which show that the preliminary phase is decisive.

We can now turn to some theoretical systems which might claim to embody an explanation of our results:

- (1) *The Gestalt school and perceptual organization*. Many of the subjects, when questioned about their experiences during the experiment, reported that the changed stimulus 'stood out' among the rest. Many psychologists, on reading of something 'standing out' from its setting, will be reminded at once of the figure-ground phenomenon, and will hold that this is a clear case for explanation by the principles of the Gestalt school. But if the phenomenon is really due to the organization of the perceptual field and to the fact that the changed stimulus, being different from the others, 'stood out', we may ask why it did not occur in the control groups, for whom the odd stimulus also contrasted with those accompanying it. Until it is possible to predict when 'standing out' will occur and when it will not, it is difficult to decide how far an account in these Gestalt terms constitutes a genuine explanation, and how far it is merely a more picturesque restatement of the fact that the changed stimulus was attended to more than the others.
- (2) *Thorndike and mental fatigue*. A psychologically unsophisticated person might, if questioned, attribute our results to the fact that the subject 'got tired' of responding to the odd, unchanged stimuli. It was evidently not a matter of *muscular* fatigue, since the same muscles were used to press similar Morse keys in response to both changed and unchanged stimuli. We are therefore faced with the problem of *mental* fatigue.

Thorndike (22), p. 5) defines 'mental work' as 'work done by the animal's connexion system'. He attributes 'mental fatigue' to the fact that 'work without rest... becomes less satisfying (1) by losing the zest of novelty, (2) by producing ennui, a certain intellectual nausea, sensory pains and even headache, and (3) by imposing certain deprivations—for instance, from physical exercise, social intercourse, or sleep' (p. 122).

It is true that Thorndike is particularly concerned with falls in the work-curves of such activities as arithmetical operations, but in so far as the unintellectual task used in our experiments involved the 'connexion-system', this theory of mental fatigue may be applicable to the weakening of the connexion between unchanged stimulus and key-pressing. However, the three factors cited by Thorndike to explain mental fatigue require explanation themselves. It is evident that the analysis needs to be carried much further.

(3) *Hull and conditioned inhibition*. When obliged to choose between several alternative theories, all of which purport to explain a given set of empirical data, a scientist's first reaction is to work out and verify the different observable phenomena that can be deduced from them. But it is not uncommon to find oneself faced with more than one theoretical structure yielding similar predictions. One is then obliged to judge between them on formal grounds: their precision and operational meaningfulness, their logical rigour and their parsimony in deriving a wide variety of phenomena from a small number of assumptions.

Of contemporary psychological procedures, the one which appears to have taken most seriously the pronouncements of modern logicians and philosophers of science on the functions and desiderata of scientific theory-building is that of Hull's behaviour theory (12, 13). Hilgard, an avowed opponent of the Hullian point of view, goes so far as to say that 'whatever its deficiencies in detail, the system which Hull proposes has set a model for other theorists to emulate' (11), p. 113). Moreover, the principal factual controversies between Hull and his chief adversaries appear to be resolving themselves in his favour (5, 14). It is therefore of particular interest to find that our present experimental results can be shown to follow from Hull's postulates as they now stand, particularly from his concept of 'conditioned inhibition', derived from observations made in Pavlov's laboratory.

We shall take in turn the four principal conclusions which our experiments seem to warrant, and we shall use the most recent version of Hull's postulates (13) to account for them, although the earlier and more familiar version (12) would yield similar conclusions (ChS. = changed stimulus, UchS. = stimulus remaining unchanged.)

(1) *That a recently changed stimulus is more likely to be responded to than one which remained unchanged and has been responded to for some time* (Experiments 1-4)

(a) When there are two or more stimuli evoking incompatible reaction-tendencies at once, the response with the greatest 'momentary effective reaction-potential' ($s\ddot{E}_R$) will alone occur (Corollary xiii).

(b) $s\ddot{E}_R$ decreases as 'inhibitory aggregate' ($\dot{I}_R$) increases (Postulates X A, X C, Corollary ix).

(c) $\dot{I}_R$ is a tendency which every response, as it recurs, generates against its own repetition. It consists of two components, 'reactive inhibition' (I_R), a temporary fatigue effect, and 'conditioned inhibition' (sI_R), a learned reluctance to respond, which becomes attached to the stimulus accompanying the cessation of the response (X A, X B, X C, Corollary ix).

Of the two components of $\dot{I}_R$, I_R concerns a particular response, and therefore, since similar responses were made to all stimuli, this cannot be responsible for our finding. On the other hand, sI_R is associated with particular stimuli, and so differences in this variable are relevant to preferences for responding to one stimulus rather than another.

Now, during the preliminary series with homogeneous stimuli (Phase I, experimental group), each response will add an increment of I_R and sI_R (X A, Corollary ix). But as long as the $s\dot{E}_R$ of all the stimuli remain above the 'reaction-threshold' (sL_R) (XIV B), all are able to evoke responses.

When ChS. is introduced, it will have had less time to acquire sI_R than UchS., and so, within the limits determined by 'behavioural oscillation' (sO_R), its $s\dot{E}_R$ will be greater more often than not. ChS. is therefore more likely to be responded to. In the control groups the odd stimulus will not have this advantage, owing to the absence of the preliminary phase.

(b) *That the effect diminishes with time* (Phases II and III, experimental group, Experiments 1-3)

(a) Reactive inhibition (I_R) is a 'negative drive' (fatigue effect) attached to a particular response and increasing with every evocation of that response (X A). In our experiments similar responses were made to both UchS. and ChS., and so we can expect I_R to increase every time a key is pressed, no matter what the stimulus.

(b) The existence of conditioned inhibition (sI_R) is deduced as a corollary (Corollary ix) from the principles governing learning in general. It may therefore be presumed that it will increase, like any learned tendency, with the number of reinforcements, according to the negatively accelerated, concave-downward growth-curve (IV). As far as sI_R is concerned, every response means a reinforcement, since its cessation relieves I_R .

(c) The increment of conditioned inhibition (ΔsI_R) occurring with each response is stated to be an increasing function of the I_R present (Corollary ix).

Now, in our experimental groups, UchS. has been responded to forty-eight times (Phase I) before ChS. first appears. UchS. will therefore have acquired a considerable amount of sI_R , and the response of key-pressing a considerable amount of I_R , by the time Phase II begins. But as ChS. comes to evoke responses during Phases II and III, it will be acquiring more and more sI_R in its turn.

Furthermore, during Phases II and III, the sI_R of ChS. will increase more rapidly than the sI_R of UchS. for two reasons:

(i) The growth curve of sI_R -accumulation for ChS. will have a steeper gradient, since it is in its early stages, than the curve for UchS., which has been continuing for some time.

(ii) The curve for ChS. will, in any case, have a steeper gradient in its early stages than the corresponding stages of the curve for UchS., since more I_R is present and sI_R will consequently accumulate more rapidly.

It follows, therefore, from the shapes of the curves, that the difference in sI_R between ChS. and UchS. will gradually diminish, and the greater probability of response to ChS. will gradually wane.

All this, of course, takes no account of the complications due to stimulus generalization of sI_R (XI), but these should not be such as to alter the trends in question.

- (3) *That the effect is not so apparent when there are several changed stimuli and only one remaining unchanged* (Experiment 1, Control Group B)

With a larger number of ChS. the sI_R for ChS. can be expected to increase more rapidly and so the difference in sI_R between ChS. and UchS. will disappear more quickly.

- (4) *That the effect is particularly strong and persistent if the changed stimulus continues to undergo changes* (Experiment 4)

When ChS. changes frequently (i.e. in Hullian language 'shifts along generalization continua'), each change will mean a decrease in 'effective inhibitory potential' sI_R (XIC) so that ChS. will not accumulate sI_R so quickly. Its power to attract attention will therefore continue for a longer time.

Attention and the perceptual response. This discussion, like the study to which this paper is a sequel (2), has so far regarded the probability of response to one stimulus rather than another as a manifestation of the direction of attention. But it is possible to treat the question more thoroughly if a suggestion put forward elsewhere (4) by the writer is accepted. This was that Hull's behaviour theory may be extended to the treatment of perception and attention by regarding perception as a stimulus-producing response ('pure stimulus act') or complex of such responses. Attention then fits in neatly as the 'momentary effective reaction-potential' of the perceptual response ($s\ddot{E}_R$).

All the above reasoning, concerning the conditions in which the overt response to ChS will have a higher momentary effective reaction-potential ($s\ddot{E}_R$) and thus be more likely to occur than the response to UchS., will imply in the same way that the perceptual response to ChS. will have a higher $s\ddot{E}_R$ in those conditions, i.e. that ChS. will receive more attention.

This makes the assumption that Hull's I_R and sI_R apply also to the perceptual response which is thought of as a neural process in the cerebral cortex. This theory of inhibition seems to fit the facts (see Berlyne (4)) and to be required to explain such phenomena as spontaneous reversals in ambiguous figures, figural after-effects and boredom (e.g. Karsten's 'satiation' experiments (15)). Prolonged perceptual responses appear to arouse a growing reluctance to continue them, which disappears partially after a rest but usually leaves a permanent residue—a similar pattern to that followed by prolonged overt responses and explained by Hull's theory (12), chaps. XV, XVI).

But I_R and sI_R , as phenomena affecting muscular and glandular responses, are stated to be due to fatigue, and it is notorious that nerve cells are not subject to fatigue apart from the brief refractory period. Is it not therefore unlikely that perceptual responses follow similar principles?

We can reduce the dubiousness of our procedure by noting first that physiologists are coming to distrust more and more the old waste-products explanation of fatigue, even for muscles. As one recent and reputable text-book puts it, 'the evidence is rather strong that such a theory cannot constitute a complete explanation' (10, p. 366). Another says, 'no one would now attempt to give so simple an explanation of the various sorts of fatigue', and attributes them rather to disturbances of equilibrium between the 'very large number of processes which interact so as to preserve the condition of homeostasis' (25, p. 16).

Gerard (8) has indicated the existence of such changes, involving a new equilibrium at a lower level of activity, in long-continued neural processes, and he has proposed the name 'equilibration' for these changes. In an elaborate discussion of the effects of monotony

cerebral activities, Hebb⁽⁹⁾ suggests that the temporary weakening of response-tendencies may be due to some breakdown in the co-ordination between cell-assemblies. On the other hand, the view that even muscular fatigue depends on a mutual interference between neural processes, including kinaesthetic impulses, is not new to experimental psychologists (see Myers⁽²¹⁾, pp. 174-5).

All these contributions point to a close affinity between nerve and muscle in the impaired performance which both show after prolonged stimulation, so that, according to Forbes⁽⁷⁾, p. 173), 'the fact that fatigue can be demonstrated in nerve tends to make the difference between nerve and muscle quantitative rather than qualitative in this respect'. Our suggestion that similar principles may underlie overt and perceptual responses is therefore not so audacious after all.

Conclusion. Our experiments were designed to illustrate the attraction of attention by change in as simple a form as possible. Their very simplicity is bound to make them seem artificial and redolent of the laboratory, like most of the human and animal investigations in which behaviour theory has to rely for its data. However, such experiments, despite their apparent remoteness from real-life situations, are necessary for the following reasons: (1) Both the hypothesis that recently changed stimuli attract attention and the predictions from the postulates of behaviour theory can be given a much more rigorously controlled verification than a more complex and realistic situation would permit.

(2) The relations between the phenomenon with which we are concerned and other aspects of behaviour can emerge much more clearly. The effect of change need not be accepted as an irreducible fact, but its analysis can be carried further: it can be shown to be a special case of more widely applicable principles, which have been found to have explanatory value for other facts also.

(3) A study of simple prototypes is a useful preliminary to the investigation of more complicated everyday phenomena. It can indicate some of the variables involved and the functional relations between them. It can equip us with a stock of concepts, principles and hypotheses, so that we can face the intricacies of complex cases with less embarrassment.

IV. SUMMARY

The tendency to attend to recently changed stimuli, which forms the subject of this paper, is distinguished from certain other determinants of attention which might also be thought of as cases of 'novelty' or 'change'.

Experiments are then described which apply the method of preferential responses to the study of this tendency and the task of demonstrating its existence. The results favour the following conclusions:

(1) That a recently changed stimulus is more likely to be responded to than one which has remained unchanged and has been responded to for some time.

(2) That the effect diminishes with time.

(3) That the effect is not so apparent when there are several changed stimuli and only one remaining unchanged.

(4) That the effect is particularly strong and persistent if the changed stimulus continues to undergo changes.

In the course of a discussion on the theoretical implications of these phenomena they are found to be deducible from the most recent version of Hull's postulates of behaviour theory. Their bearing on a recent suggestion for integrating the psychology of attention and perception with Hull's behaviour theory is then considered.

REFERENCES

- (1) ADRIAN, E. D. (1947). *The Physical Background of Perception*. Oxford University Press.
- (2) BERLYNE, D. E. (1950). Stimulus intensity and attention in relation to learning theory. *Quart. J. Exp. Psychol.* II, 71-5.
- (3) BERLYNE, D. E. (1950). Novelty and curiosity as determinants of exploratory behaviour. *Brit. J. Psychol.* XLI, 68-80.
- (4) BERLYNE, D. E. (1951). Attention, perception and behaviour theory. *Psychol. Rev.* LVIII, 137-46.
- (5) BLUM, R. A. & BLUM, J. S. (1949). Factual issues in the 'continuity' controversy. *Psychol. Rev.* LVI, 33-50.
- (6) DESAI, M. M. (1939). Surprise. *Brit. J. Psychol. Monogr. Supp.* XXII.
- (7) FORBES, A. (1934). The mechanism of reaction (in MURCHISON, C. (ed.). *A Handbook of General Experimental Psychology*. Worcester, Mass.: Clark University Press).
- (8) GERARD, R. W. (1927). Studies on nerve metabolism. II. Respiration in oxygen and nitrogen. *Amer. J. Physiol.* LXXXII, 381-404.
- (9) HEBB, D. O. (1949). *The Organization of Behavior*. New York: Wiley.
- (10) HEILBRUNN, L. V. (1943). *An Outline of General Physiology*, 2nd ed. Philadelphia: Saunders.
- (11) HILGARD, E. R. (1948). *Theories of Learning*. New York: Appleton-Century.
- (12) HULL, C. L. (1943). *Principles of Behavior*. New York: Appleton-Century.
- (13) HULL, C. L. (1950). Behavior postulates and corollaries—1949. *Psychol. Rev.* LVII, 173-80.
- (14) JENKINS, W. O. & STANLEY, J. C. (1950). Partial reinforcement: a review and critique. *Psychol. Bull.* XLVII, 193-234.
- (15) KARSTEN, A. (1928). Psychische Sättigung. *Psychol. Forsch.* x, 142-254.
- (16) MEEHL, P. E. & MACCORQUODALE, K. (1948). A further study of latent learning in the T-maze. *J. Comp. Physiol. Psychol.* XLI, 372-96.
- (17) MILLER, N. E. (1941). An experimental investigation of acquired drives. *Psychol. Bull.* XXXVIII, 534-41.
- (18) MILLER, N. E. & DOLLARD, J. (1941). *Social Learning and Imitation*. London: Kegan Paul, 1945.
- (19) MOWRER, O. H. & LAMOREAUX, R. R. (1942). Avoidance conditioning and signal duration—a study of secondary motivation and reward. *Psychol. Monogr.* LIV, 5. (Reprinted in (20).)
- (20) MOWRER, O. H. (1950). *Learning Theory and Personality Dynamics*. New York: Ronald.
- (21) MYERS, C. S. (1931). *A Text-Book of Experimental Psychology*, I, 3rd ed. Cambridge University Press.
- (22) THORNDIKE, E. L. (1914). *Mental Work and Fatigue and Individual Differences and their Causes* (vol. III of *Educational Psychology*). New York: Teachers' College, Columbia University.
- (23) TITCHENER, E. B. (1901). *Experimental Psychology*, vol. I, part II. New York: Macmillan.
- (24) WILCOCKS, R. W. (1928). The effect of an unexpected heterogeneity on attention. *J. Gen. Psychol.* I, 266-319.
- (25) WINTON, F. R. & BAYLISS, L. E. (1948). *Human Physiology*, 3rd ed. London: Churchill.

(Manuscript received 27 February 1951)

THE IDENTIFICATION OF MORSE SIGNALS WHEN HEARD AGAINST A CONTINUOUS TONE

By R. C. BROWNE

*From the Nuffield Department of Industrial Health, University of Durham
(King's College, Newcastle upon Tyne)*

*The problem (pp. 279-280). II. Apparatus and methods (pp. 280-281). III. Results (pp. 281-286).
IV. Discussion (pp. 286-287). V. Summary (p. 287). References (p. 287).*

I. THE PROBLEM

Turning the machine the wrong way in response to hearing a given signal was the commonest mistake made during the recent war in landing an aircraft by the beam-approach procedure (Browne⁽²⁾, 1942). A natural question to ask is how important, in making this mistake more or less likely, is the intelligibility of the morse signal which is sent to the pilot? Interest was lent to this problem in that the United States Army Air Force used a different method which sent the letters A and N to the pilot, instead of E and T as in the British system (Fig. 1). Moreover it was necessary to decide between these two methods with a view to eventual standardization.

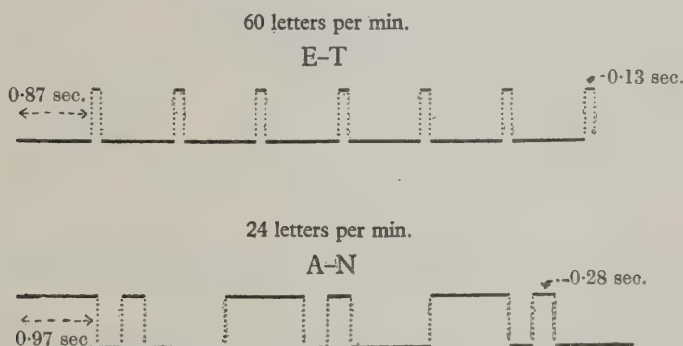


Fig. 1. A comparison of the E-T and A-N aural signals.

The general problem can be considered under the following subsidiary headings:

- (1) Are the morse letters A and N more easily identifiable than the letters E and T?
- (2) Can one pair of letters be detected at a lower contrast level, or keying depth, than the other?

Both morse letter and background sound have a frequency of 1150 cyc./sec. and are differentiated from each other only by volume. The keying depth is the volume in decibels of the morse letter above the volume of the continuous background sound which is simultaneously transmitted.

- (3) Is there a relationship between the volume of the morse signal letter and the keying depth at which it can first be detected?

(4) What is the relation between the speed of sending and the accuracy of letter identification?

The engineers were anxious to know this relationship, because they felt that they ought to send the signals as fast as possible to secure a quick response, but they gave less attention to the danger of a quick response which might be wrong. Moreover, it has been pointed out by Bartlett⁽¹⁾ (1943) that in auditory display systems there is a tendency to cut down the time of exposure below economic levels.

II. APPARATUS AND METHODS

The apparatus (Fig. 2) was made to reproduce the various signals which were to be studied. It is 1 cu.ft. in size, portable, and can be plugged into any electric-light socket. It was

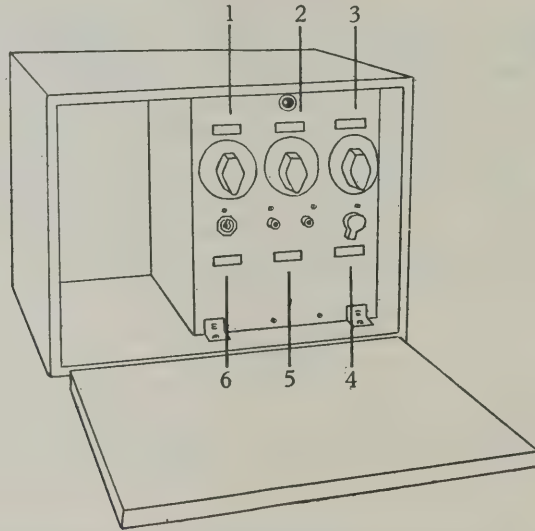


Fig. 2. Diagram of apparatus for synthesizing the audible signals. (1) Speed of sending; (2) volume of signal. (3) keying depth; (4) selector switch for E-T or A-N; (5) telephone socket; (6) main switch.

designed so that the variables in the problem can be accurately controlled and was made by the Radio Department, Royal Aircraft Establishment, Farnborough.

The power supply to the mechanism is through a valve rectifier which feeds the current into a low-frequency oscillator and amplifier which produces the basic tone. The volume is controlled by a knob, calibrated to show the potential difference across the telephone terminals. The square of this difference is proportional to the volume in decibels above audible threshold. The output of the oscillator goes through motor-driven interrupter cams, which automatically convert it into the right combination of dots and dashes to form the letters E and T, or A and N. The speed of this motor is controlled by a rheostat which is calibrated to show the relation between the actual sending speed and the normal speed of the system concerned. Finally, there is a keying depth control, so that the volume of the morse signal can be made to stand out, up to a maximum of 18%, above the volume of the background tone. Two pairs of plug-in telephones are fitted, one to be worn by the subject and one by the experimenter.

The subjects of this experiment were forty air-crew cadets chosen at random, who satisfied the selection requirements for pilots and navigators. They had been trained in the morse code, but had not yet been taught to fly. A full explanation was given to them of what the experiment was about, and they were told the four possible letters which they might hear and also that a letter might be heard twice. The test was done twice on the subjects, one-half of whom were sent E and T first and the other half A and N. The subject sat on one side of the test apparatus and put on one of the pairs of telephones, while on the other side sat the tester who wore the other pair and had by him a stop-watch. The subjects called out the letters as soon as they identified them. The signals were sent at speeds which were normal for the two systems then in use, and which provided the *ad hoc* problem. The speeds were 60 letters per min. for E and T, and 24 letters per min. for A and N. The test set was switched on, and the keying depth was started at zero for each man and then gradually increased at a constant speed of 1% in 2 sec. until the subject named the letter rightly or wrongly, when the signals were switched off, and the time and keying depth noted. The rate of increase of keying depth was, therefore, constant for each of the four letters and from subject to subject. The volume of the signals produced by the apparatus was measured by a sound-level meter, and the relation between the calibrations of the volume control and the volume in decibels above audible threshold is shown in Table 1.

Table 1. *The relation of the volume control calibration to the sound intensity in the telephones. Each figure is the average for four telephones*

Volume control calibration	Volume in telephones db. above threshold
1	16.2
2	22.3
3	27.0
4	30.3
5	32.3
10	37.5
15	40.4
20	42.5

III. RESULTS

(1) *The relative identifiability of the four letters E-T and A-N and the effects of different keying depths and signal volumes*

Table 2 shows the number and proportion of the forty subjects who identified each letter pair rightly or wrongly, and it demonstrates that with the letters A and N the proportion of subjects who make mistakes is $25 \pm 10.1\%$ less than with the letters E and T. In view of the magnitude of this difference and the reproducibility of the result in both groups of cadets, this finding is not likely to be due to chance. A further group of forty subjects, chosen at random, was now taken to avoid the influence of practice, and it was used to

Table 2. *The number and proportion of two groups of 20 subjects who did and did not make mistakes in identifying the letters*

Keying system ...	E-T				A-N			
	Without mistakes		With mistakes		Without mistakes		With mistakes	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Experiment								
Absolute no. of subjects	11	11	9	9	18	14	2	6
Percentage of total no.	55	55	45	45	90	70	10	30

determine whether increasing the volume of the signal four times made any difference to the speed (or keying depth) and accuracy of the identification of the letters. It was found, however (Table 3), that this increase in volume from 16.2 to 22.3 db., which represents a four times increase in intensity, made no difference to the speed of identification of the letters. In short, therefore, there is nothing to be gained, either in speed of reception or in depth of keying, by increasing the volume of the signal by this amount. There is, on the other hand, a slight but consistent increase in the number of mistakes made at the higher volume, since they are 8 and 9 for the two volumes for E, 6 and 7 for T, 0 and 2 for A, and 2 and 4 for N. The higher figure is in each case to be correlated with the higher volume, which suggests the tentative conclusion that, on the one hand, there is nothing to be gained in speed of reception by such an increase in volume, and on the other, that there is possibly something to be lost in accuracy.

Table 3. *The effect of increasing the signal volume four times upon the speed or keying depth and accuracy of identification of the letters*

Letter sent ...	E		T		A		N	
Times increase in volume	1	4	1	4	1	4	1	4
Keying depth percentage above volume of beam background	8.6	7.6	7.8	5.9	9.7	7.9	7.9	8.6
No. of mistakes	8	9	6	7	0	2	2	4
Total number	30				8			

Table 4 carries the inquiry a step further forward, and shows the number of subjects out of this group who made 0, 1 and 2 mistakes in each letter pair. It confirms the previous finding by demonstrating that as many as thirty-two men, or 80%, make no mistakes when they are identifying A and N, but only twenty-one, or 52.5%, avoid error when identifying E and T. Moreover, no subject makes two mistakes with A and N, but eleven men, or 27.5%, do so with E and T. It appears, therefore, that more people correctly identify the morse letters A and N than E and T, and that few, if any, get them both wrong.

Table 4. *The number and proportion of subjects who made 0, 1 and 2 mistakes with each pair of letters*

Keying system ...	E-T			A-N		
No. of mistakes	0	1	2	0	1	2
No. of subjects	21	8	11	32	8	0
Percentage of total no. of subjects	52.5	20.0	27.5	80.0	20.0	0

The individual differences in the times taken to name rightly each of the four letters are shown in Table 5. For E and T the distribution is skewed, with a peak at 34% correct identifications at a keying depth of 5-6%, or an identification time of from 10 to 12 sec. The A and N distribution is, however, much more scattered, with two peaks of 19.4% correct identifications at 12-14 sec. and 18% at 20-22 sec. Since, however, *correct* identifications only are being considered, and the sending speed of E and T is 60 letters per min., whereas that of A and N is only 24, this probably means that the letters E and T are either correctly named fairly soon or they are not. It may, however, take rather longer to identify A and N, which is eventually done correctly more often, accuracy being bought at the expense of speed.

Table 5. *The individual differences in the speed and accuracy of letter identification*

Keying depth (%)	Time taken to name letter right in sec.	No. of times letter was identified rightly					
		E T		Totals E + T		A N	
				No.	%		
0-4	0-8	—	3	3	6	—	1
4-5	8-10	4	1	5	10	—	7
5-6	10-12	8	9	17	34	7	3
6-7	12-14	1	6	7	14	10	4
7-8	14-16	2	3	5	10	5	1
8-9	16-18	1	—	1	2	2	1
9-10	18-20	2	1	3	6	3	3
10-11	20-22	—	2	2	4	5	8
11-12	22-24	—	—	—	—	—	1
12-13	24-26	3	2	5	10	1	1
13-14	26-28	1	—	1	2	1	1
14-15	28-30	—	—	—	—	1	—
15-16	30-32	—	—	—	—	3	—
16 and over	32 and over	1	—	1	2	1	2
Totals		23	27	50	100	39	33
						72	100.0

(2) *The relationship between the speed of sending and the speed and accuracy of identifying the letters*

The effect of changing the speed of sending upon the accuracy of identification was studied by taking 160 air-crew cadets of a type similar to those already described, and by allocating them at random into four groups, A, B, C and D, each containing forty men to be tested with a different speed of transmission. It was necessary to use separate groups of subjects for every speed, because if the same group had been used for all speeds the effect of improvement with practice might have seriously blurred the results. Each test was repeated three times on a given man, who was told that he would only hear two letters, E and T or A and N, so that there were only two possibilities from which to choose instead of four as in the previous experiment. In that experiment the problem was the intrinsic identifiability of the four different letters, whereas the problem in the present experiment was to compare and study the E-T and A-N systems as they stand. A random method was used to decide whether any given subject received an E or a T, or an A or an N first, as the adoption of any regularly recurring system might have been discovered in time by the subjects themselves. The original speeds of sending were 60 letters per min. for the English E-T system and 24 letters per min. for the American, or A-N method. The speeds used in this experiment were 16, 30, 60 and 120 letters per min. for E and T, and 24, 48, 72 and 96 letters per min. for A and N. These are approximate multiples of the original speeds, and it was hoped that the differences between them would give fairly clear-cut results.

Table 6 and Fig. 3 show the relation between the speed of sending and the proportion of men who correctly identify the letters. A rise in the speed of sending E-T from 16 to 30 letters per min. increases the proportion of subjects who name them rightly from 45 to 65 %, with a difference of 20 ± 6.3 % which suggests that the improvement is not likely to be due to chance. The proportion giving right answers falls, however, from 65 to 60, and finally to 51 % as the sending speed rises from 30 to 60, and again to 120 words per min. The fall in accuracy of reception between the sending speeds of between 30 to 60 letters per min. may be due to chance, but it is confirmed by a further fall when the speed rises to 120 letters per min. Between speeds of 60 and 120 each of the three groups of subjects is consistently

Table 6. *The number and proportion of subjects who identified the letters of the two keying systems rightly and wrongly. Four different speeds of transmission were used on each system*

Keying speed letters per min.	Subject's group letter	No. of test	Keying system							
			E-T				A-N			
			Without mistakes		With mistakes		Without mistakes		With mistakes	
			No.	%	No.	%	No.	%	No.	%
16	D	1	13	32.5	27	67.5	—	—	—	—
		2	19	47.5	21	52.5	—	—	—	—
		3	22	55.0	18	45.0	—	—	—	—
		Mean	18	45.0	22	55.0	—	—	—	—
24	A	1	—	—	—	—	31	77.5	9	22.5
		2	—	—	—	—	30	75.0	10	25.0
		3	—	—	—	—	33	82.5	7	17.5
		Mean	—	—	—	—	31	78.0	9	22.0
30	C	1	27	67.5	13	32.5	—	—	—	—
		2	27	67.5	13	32.5	—	—	—	—
		3	24	60.0	16	40.0	—	—	—	—
		Mean	26	65.0	14	35.0	—	—	—	—
48	B	1	—	—	—	—	30	75.0	10	25.0
		2	—	—	—	—	30	75.0	10	25.0
		3	—	—	—	—	29	72.5	11	27.5
		Mean	—	—	—	—	30	75.0	10	25.0
60	A	1	23	57.5	17	42.5	—	—	—	—
		2	23	57.5	17	42.5	—	—	—	—
		3	26	65.0	14	35.0	—	—	—	—
		Mean	24	60.0	16	40.0	—	—	—	—
72	D	1	—	—	—	—	25	62.5	15	37.5
		2	—	—	—	—	34	85.0	6	15.0
		3	—	—	—	—	28	70.0	12	30.0
		Mean	—	—	—	—	29	73.0	11	27.0
96	C	1	—	—	—	—	25	62.5	15	37.5
		2	—	—	—	—	29	72.5	11	27.5
		3	—	—	—	—	27	67.5	13	32.5
		Mean	—	—	—	—	27	68.0	13	32.0
120	B	1	19	47.5	21	52.5	—	—	—	—
		2	21	52.5	19	47.5	—	—	—	—
		3	21	52.5	19	47.5	—	—	—	—
		Mean	20	51.0	20	49.0	—	—	—	—

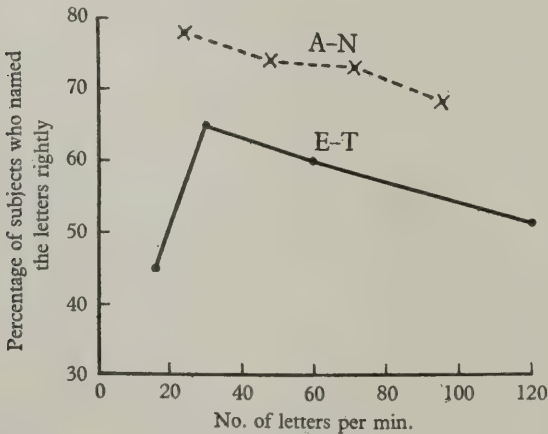


Fig. 3. The accuracy of letter identification at various sending speeds.

worse, the comparable proportions who did not make mistakes at the two speeds being 57.7 and 47.5, 57.5 and 52.5, and finally 65.0 and 52.5%. These findings suggest, therefore, that the E-T keying system is most accurately received between the sending speeds of 30 and 60 letters per min.

The results with the A-N method contrast with the above, since a slight but consistent fall in the proportion of subjects who named the letters rightly from 78% through 75 and 73 to 68% was found as the speed of sending *rose* from 24 letters per min. through 48 and 72 to 96 letters per min. These differences are not great enough to be technically significant, but their systematic consistency suggests a true fall in accuracy of reception with increase in speed of sending with this keying method. Moreover, inspection of the results of the subject groups A, B, C and D in Table 6 shows that no one group is intrinsically better or worse than another.

The individual differences in the times needed to identify the letters at different sending speeds are shown in Tables 7 and 8. With the letters E and T most identifications are made after listening for 8-20 sec. (Table 7) when the sending speed is 16 letters per min., but

Table 7. *The individual differences in the speed and accuracy of identifying E and T at four sending speeds*

No. and proportion of correct identifications of E or T
at four speeds of sending in letters per min.

Seconds taken to identify letter	16		30		60		120	
	No.	%	No.	%	No.	%	No.	%
0-6	3	2.1	2	1.1	1	0.5	7	4.4
6-8	3	2.1	3	1.6	26	13.6	27	16.9
8-10	14	9.7	9	4.9	21	11.0	24	15.0
10-12	17	11.7	20	10.9	45	23.6	34	21.3
12-14	17	11.7	26	14.2	23	12.7	23	14.4
14-16	12	8.5	19	10.4	15	8.3	17	10.6
16-18	20	13.8	18	9.9	13	6.8	11	6.9
18-20	14	9.7	18	9.9	13	6.8	9	5.6
20-22	8	5.5	24	13.1	12	6.3	3	1.9
22-24	7	4.8	14	7.7	5	2.6	3	1.9
24-26	9	6.2	12	6.6	1	0.5	1	0.6
26-28	4	2.8	5	2.7	—	—	1	0.6
28-30	2	1.4	10	5.5	1	0.5	—	—
30-32	7	4.8	3	1.6	1	0.5	—	—
32-40	8	5.5	—	—	4	2.1	—	—

when the speed rises to 30 there is a paradoxical lengthening of the identification times with an increase in their scatter. At sending speeds of 60 and 120 letters per min. more of the individual identifications are distributed at the rapid end of the scale with a smaller overall scatter. With A and N, at a sending speed of 24 letters per min. (Table 8) most of the identification times lie between 8 and 18 sec., but as the speed rises to 48 and 72 letters per min., so the individual observations tend to be more quickly made. At a sending speed of 96 letters per min., however, the individual differences in the identification times tend to be rather more marked, which suggests that at this speed intelligibility is becoming difficult.

Table 9 makes a general comparison of the two systems. Considering first the E-T method and taking each keying speed in turn, it can be seen that a sending speed of 16 letters per min. is worst, because it has the lowest proportions of subjects making correct identifications (45%), the longest time to identify the signals (17.7 sec.), and the greatest

scatter around this value (7.3 sec.). Similarly, it can be seen that 60 letters per min. is probably the best all round sending speed, since it has a slightly greater proportion (60%) of men who made no mistakes, and its mean time for correct identification is 13.3 sec., which is 4.4 sec. quicker than its nearest rival which has a sending speed of 30 letters per min. Inspection of the figures for the A-N method shows little to choose between the four

Table 8. *The individual differences in the speed and accuracy of identifying A and N at four sending speeds*

Seconds taken to identify letter	No. and proportion of correct identifications of A or N at four speeds of sending in letters per min.							
	24		48		72		96	
	No.	%	No.	%	No.	%	No.	%
0-6	—	—	8	4.0	9	4.5	8	4.1
6-8	4	2.0	32	15.8	45	22.7	28	14.4
8-10	26	12.7	50	24.8	42	21.2	31	15.9
10-12	70	34.3	39	19.3	38	19.2	31	15.9
12-14	34	16.7	22	10.9	19	9.6	26	13.3
14-16	19	9.3	20	9.9	18	9.1	31	15.9
16-18	22	10.8	11	5.4	12	6.1	21	10.8
18-20	9	4.4	6	3.0	3	1.5	8	4.1
20-22	9	4.4	3	1.5	7	3.5	2	1.0
22-24	2	1.0	4	2.0	3	1.5	5	2.6
24-26	4	2.0	1	0.5	—	—	—	—
26-28	3	1.5	—	—	—	—	1	0.5
28-30	—	—	1	0.5	1	0.5	1	0.5
30-32	1	0.5	1	0.5	1	0.5	—	—
32-40	1	0.5	3	1.5	—	—	2	1.0

Table 9. *The relationship of the keying or sending speed to the speed and accuracy of reception for the two systems*

	Keying system							
	E-T				A-N			
Keying speed, letters per min.	16	30	60	120	24	48	72	96
Percentage with correct identifications	45	65	60	51	78	75	73	68
Mean time for correct identification (sec.)	17.7	17.7	13.3	11.9	13.0	11.9	11.1	12.5
Standard deviation (sec.)	7.3	6.0	5.6	4.4	4.6	5.7	4.7	5.2

different sending or keying speeds. Perhaps, if anything, a speed of 24 letters per min. has the advantage, but as an aural indication system it is considerably better, since the proportion of subjects with no mistakes is higher, and the mean identification time and the standard deviation of the results are lower than the corresponding findings for the E-T method.

IV. DISCUSSION

Vernon⁽³⁾ (1934) has suggested that the figure and ground principle holds good in auditory perception, as it does in visual. Moreover, he pointed out that in normal listening the figure may contrast with the ground most importantly in pitch, intensity, movement, timing and dynamics. In the present problem, however, the contrast is restricted, in hearing E or T, to only intensity and timing. The subject, therefore, either in the aircraft or in the laboratory, is faced with deciding whether he is hearing a 'figure' of morse dots or of dashes against a continuous 'ground' of the same pitch but slightly different intensity. It was a general complaint of air pilots that this was very difficult, and quantitative support was lent to their dissatisfaction by a report (Browne⁽²⁾, 1942) which showed that it

was common for a pilot to turn the wrong way on hearing a given morse signal. Moreover, the present study shows that 40% of subjects mistook dots for dashes at the usual keying speed of 60 per min. It seems reasonable to explain the greater accuracy of response to hearing the morse letters A (·—) and N (—·) by the fact that here there is an additional element of contrast in the dynamics of the system, which is absent when E (·) and T (—) are compared to their background.

V. SUMMARY

1. A comparison has been made of the speed and accuracy of identification of the morse letters A or N and E or T, transmitted at an audible frequency of 1150 cyc./sec., and contrasting in intensity with a background tone of the same frequency.
2. Under these conditions the letters A and N are identified with greater speed and accuracy than are E and T.
3. Increasing the loudness of the signals from 16 to 22 db. above the hearing threshold has little effect upon the speed of identification.
4. Increasing the speed of sending of E and T from 20 to 40 letters per min. increased the accuracy of identification, but more rapid sending decreased it. But increasing the speed of sending of A and N from 24 to 96 letters per min. decreased the accuracy of identification.
5. A tentative explanation of the results in terms of the figure and ground principle is made.

Acknowledgements are due to the Medical Directorate, Air Ministry, for permission to publish this paper.

REFERENCES

- (1) BARTLETT, F. C. (1943). *Flying Personnel Research Committee Report*, no. 565.
- (2) BROWNE, R. C. (1942). *Flying Personnel Research Committee Report*, no. 507.
- (3) VERNON, P. E. (1934). Auditory perception. 1. The Gestalt approach. *Brit. J. Psychol.* xxv, 123.

(Manuscript received 11 January 1951)

VISUAL SIZE CONSTANCY IN THE CAT

By RALPH GUNTER, *Vision Research Unit, Medical Research Council,
Institute of Ophthalmology, London*

I. *Introduction* (pp. 288–289). II. *Size constancy in binocular vision* (pp. 289–291): (i) *Training experiments and results* (p. 289); (ii) *Test experiments and results* (pp. 289–291); (iii) *Discussion and conclusions* (p. 291). III. *Size constancy in monocular vision* (pp. 291–292): (i) *Preliminary training* (p. 291); (ii) *Experimental method* (p. 291); (iii) *Test experiments* (pp. 291–292); (iv) *Results and discussion* (p. 292). IV. *Summary* (p. 293). *References* (p. 293).

I. INTRODUCTION

When two objects viewed are identical in form, and differ only in their linear dimensions, then it can be assumed that the objects throw images on the observer's retina which stand in a similar relative size relation as the actual objects in space, i.e. the image of the larger object will be large and the image of the small object will be correspondingly smaller.

If the physically larger of the two objects be moved away such a distance that its image on the retina now is smaller in size than the one given by the physically smaller object near the observer, the relative size of the retinal image reverses that of the objects, and the physically larger object which is giving the smaller image should be judged smaller than the physically smaller object whose image is relatively larger. However, such judgements—out of harmony with objective reality—are not normally made owing to the fact that we are not guided in our judgement by the size of the retinal image only. We, in fact, make use of secondary cues—both visual and non-visual—which help us in our judgement to approximate the size of the real object. Martin & Pickford (9), investigating some of the conditions influencing relations of apparent size in human observers, find visual distance cues that diminish with increase in veiling glare, leading to a reduction of phenomenal regression, so that the apparent size relations of the two stimulus objects more and more nearly approach to reach a situation to be expected on the inverse square law.

These problems raise the interesting question of the difference between stimulus and sensation, and throw some doubt on the validity of cortical point to point representation from the retina. Many of the problems of 'personal space', depth perception and constancy have been investigated in human subjects, and a masterly discussion of this complex problem is given in a recent publication by Joynson (4). However, comparatively little is reported in the literature of how animals will deal with these problems.

Of all the constancy problems the question of colour constancy has been most widely studied. Köhler (7) finds colour constancy in chimpanzee and fowl. He rejects the explanation of colour constancy as being an individual acquisition, and attempts a nativistic interpretation. Katz & Révész (5) using hens, Klüver (6) on the macaque, and Locke (8) on the rhesus monkey, all confirm Köhler's findings. Even fish were found by Burkamp (1) to exhibit the phenomenon of colour constancy.

As far as constancy of size is concerned, few experiments are cited in the literature. Köhler (7) proved its existence in monkeys, and Götz (2) summarizes his results on the domestic fowl by stating that 'in spite of the inability to see stereognostically, hens reveal

behaviour which allows us to conclude that they have an amazingly highly developed size constancy'. M. Hertz (3) comes to essentially the same conclusions in her work on bees. The purpose of the present investigation was to ascertain—under controlled conditions of illumination—to what extent constancy of size could be found in the cat.

II. SIZE CONSTANCY IN BINOCULAR VISION

(i) *Training experiments and results*

Under ordinary daylight conditions in the laboratory, six cats (four males, two females, all about one year old) were trained to go for their food to a large cross 3 in. in length and having a crossbar of 1.5 in. (the positive stimulus) and to avoid a cross 1½ in. in length having

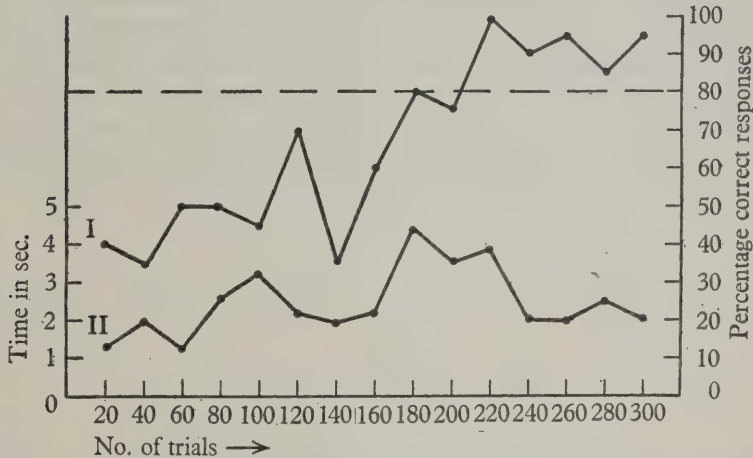


Fig. 1. Learning the absolute size discrimination. Average of six cats. Curve I: each point represents the percentage of correct responses in 20 trials. Curve II: each point represents the average running time for 20 trials.

crossbar of 0.75 in. (the negative stimulus). The lines forming the two crosses were of equal thickness. Food was put behind each door, but could only be obtained via the door bearing the positive stimulus.

The stimuli were cut out of black paper which covered a 3 × 3 in. glass panel in the doors hinged at either arm of an open Y-shaped discrimination box. Two light sources, well shielded, illuminated the stimuli from the rear with white light of equal intensity. The positive and negative stimuli were interchanged in random order for each run, so that habits of alternation and position could be controlled. Under these conditions a positive discrimination was set up in all animals in 160–200 trials. (No punishment was given for the wrong response.) The learning curves are given in Fig. 1.

(ii) *Test experiments and results*

The positive and objectively larger stimulus was then moved three times its distance from the choice point so that it now subtended the smaller visual angle—throwing an image on the cat's retina which was smaller than that of the negative stimulus (i.e. the objectively smaller cross). These tests were interspersed along with the responses to the absolute size discrimination to which the animals had been trained. This insured the

randomization of position of the stimuli and prevented the interference of the experiment with the test. Table 1 shows an example of this procedure.

Table 1. *Sample of random order used*

Run		Result	Running time (sec.)
1	Both stimuli in the same plane	+	2.0
2	Both stimuli in the same plane	+	2.5
3	(T) Objectively larger stimulus	+	6.5
4	Both stimuli in the same plane	+	3.0
5	Objectively smaller stimulus moved	-	9.0

and so on for forty runs.

With this arrangement the experiments were continued, and twenty test data were collected from forty runs. The first effect to be noted was an increase in the average running time of as much again as that recorded under the previous conditions. The animals sat at the choice point carrying out 'displacement reactions', and the percentage of the average correct responses of all animals fell below 70% in the first eighty responses to the test stimuli. In no case did the response fall below an average of 40% (see Fig. 2). While this result shows an inability to discriminate on the part of the animal, it also shows an absence of any consistent response to the negative stimulus which, under these conditions, subtended a larger visual angle.

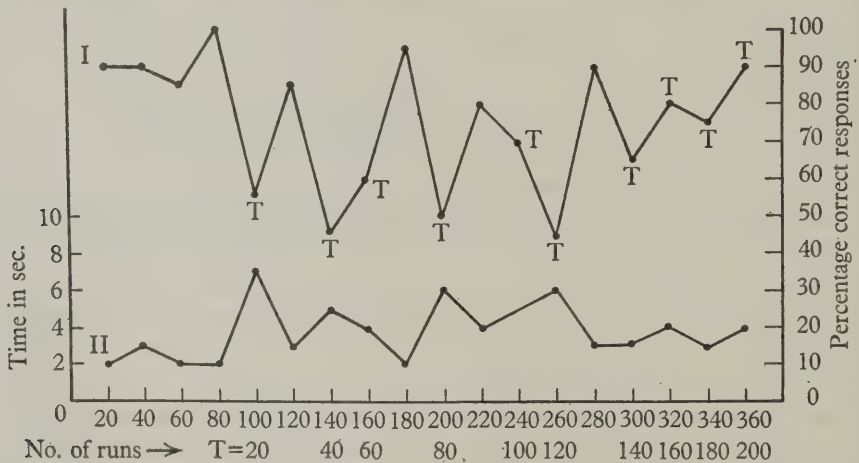


Fig. 2. Tests interposed with responses to learned discrimination. Curve I: points represent the percentage of correct responses in 20 trials. Response to absolute discrimination. T represents percentage of correct responses in 20 trials (response to test situation, i.e. stimuli subtending different visual angles). Curve II: this curve gives the mean time for 20 trials.

With this procedure and continued training all animals learned the discrimination and reached the 80% level in 160–210 runs to the test situation (see Fig. 2). This response remained consistent for all subsequent trials.

Four fresh animals of similar age but with no previous training were now introduced into the test situation to see whether a differential discrimination could be set up under these conditions.

The fresh animals were trained to react positively to the objectively larger cross, which was randomly placed farther away to subtend a smaller visual angle than the negative

stimulus cross. Forty massed trials were given daily, and the 80% level of correct response was reached by all animals only after 810-930 runs, as compared with 260-310 trials in which the original group with previous training to the objective situation learned the test problem, where both the positive and negative stimulus were situated in the same plane. At each stage of the experiment various control trials were carried out, in which the intensity of illumination of both stimuli was changed at random so that differences of response due to intensity can be said to have been ruled out.

(iii) *Discussion and conclusions*

Although further quantitative experiments are necessary for a definite conclusion, the evidence adduced above is not inconsistent with the view that even under controlled conditions size constancy seems to exist in the cat, though it has to be brought out by extensive training. However, the essential cause for the increase in running time in the new general situation (objectively larger stimulus subtending a relatively smaller visual angle) which was found in those animals which had been trained to a situation where both stimuli were in the same plane, may be sought in the general adjustment which had to be made by the animals when suddenly confronted with this new task.

III. SIZE CONSTANCY IN MONOCULAR VISION

Having adduced some evidence pointing to the existence of size constancy in the cat when binocular vision was used, the writer now proceeded to test the effect of monocular vision in this discrimination.

(i) *Preliminary training*

The same original six animals were used as subjects, and all conditions were exactly like those employed in the binocular experiment.

As a preliminary training all animals were given 100 trials using binocular vision, and a consistent correct response of between 90 and 100% to the objectively larger stimulus was given in all cases, irrespective of whether both stimuli subtended the same visual angle or not.

(ii) *Experimental method*

To test the effect of monocular vision in this discrimination, each animal had its left eye bandaged (so as to be light-tight) by means of various layers of plaster and black bandages. The bandages were applied immediately before the experiment, and the animals were not allowed to use monocular vision except during the crucial test. For the first 100 trials both the objectively larger and the objectively smaller stimulus cross were presented in the same plane, and it was interesting to observe how, after preliminary adjustments and extensive movement of the head to either side, all animals relearned the discrimination for monocular conditions in an average of 45 trials; only one animal needed sixty runs to give a consistent 80% level of correct response.

(iii) *Test experiments*

After 100 additional trials with all animals giving consistent correct responses of 80% or above where both stimuli were in the same visual plane, trials were now made in which the objectively larger stimulus cross was moved farther away to subtend a smaller visual angle than the objectively smaller stimulus cross. These test runs were intermixed in random

order with the ordinary training runs, so that the animal had to be run twenty times to give ten responses in the crucial test situation.

(iv) *Results and discussion*

From each animal 100 responses were collected in this way, all of which remained completely random. The running time under these conditions became double that of the normal one, and excessive head movements on the part of the animal were observed. The fact that of all the animals which mastered the discrimination with binocular vision, none could do it when monocular vision was employed and did not learn it after 980 trials, is probably due to the fact that the animals had no previous experience in the use of monocular vision, or that stereoscopic vision was eliminated.

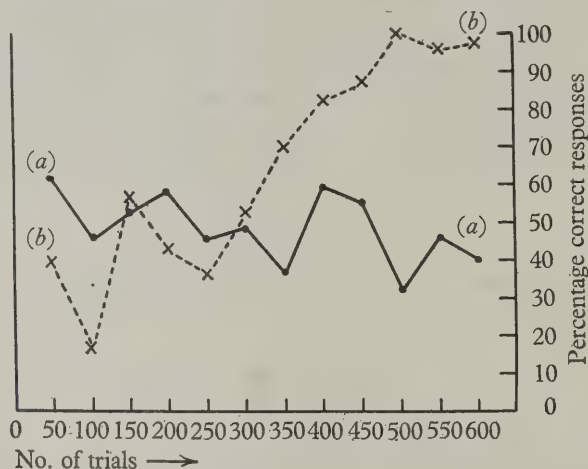


Fig. 3. Effect of previous monocular experience on learning of size constancy. Curve (a): animals using monocular vision only for test experiment. Curve (b): animals having used monocular vision for approx. 8 months prior to test experiment. Value for running time not given. Each point represents average result of 50 trials. Each cross represents average result of 50 trials.

This assumption was checked by introducing into the test situation two animals which—owing to an accident—had lost the sight of one eye approximately eight months prior to the experiment. These animals had had to rely on monocular vision in their daily habits. They were given no training in the absolute discrimination, but were immediately trained to the test experiment and gave a consistent 80% correct response to the objectively larger stimulus cross, irrespective of its position, in an average of 380 trials (as compared with 160 to 210 trials which were necessary to establish this discrimination in the group with previous training to the objective situation). Fig. 3 shows the effect of previous monocular experience on the learning of this problem.

From the above results it can be concluded that in the cat, size constancy in monocular vision is a function which has slowly to be acquired by the animal, but that with sufficient training it can be brought out for both binocular as well as monocular vision. However, the increased learning time required for the crucial size constancy reaction may simply have been necessary for adaptation to the new situation.

IV. SUMMARY

The problem of size constancy has been investigated in cats by training the animals positively to the larger of two stimulus crosses.

The animals were then tested when this positive stimulus was moved to subtend a visual angle smaller than that of the objectively smaller stimulus. Both monocular and binocular vision were employed, and it was found that

(1) Using binocular vision, size constancy seems to exist for the cat, but has to be brought out by further training.

(2) Size constancy could not be brought out, even after prolonged training, in animals which—having been trained under conditions of binocular vision—were using monocular vision during the experimental session only.

(3) Animals which had considerable experience with monocular vision prior to the experiment showed the phenomenon of size constancy, but needed almost twice as many trials as animals which used binocular vision for this reaction.

The experiments reported in this paper were carried out at the Zoology Department, Imperial College of Science, and the writer wishes to thank Asst.-Prof. H. R. Hewer for the laboratory facilities placed at his disposal.

The writer is greatly indebted to his Chief, Prof. H. Hartridge, F.R.S., Dr L. C. Thomson and Dr R. W. Pickford for their advice and criticism in the preparation of this paper.

REFERENCES

- (1) BURKAMP (1923). Versuche über das Farbenwiedererkennen der Fische. *Z. Sinnesphysiol.* LV.
- (2) GÖTZ, W. (1926). Zum Problem der Sehgrößenkonstanz beim Haushuhn. *Z. Psychol.* XCIX.
- (3) HERTZ, M. (1933). Über figurale Intensitäten. *Biol. Zbl.* vol. 53.
- (4) JOYNSON, R. B. (1949). The problem of size and distance. *Quart. J. Exp. Psychol.* 1, pt. 3.
- (5) KATZ, D. & RÉVÉSZ, G. (1921). Experimentelle Studien zur vergleichenden Psychologie. *Z. angew. Psychol.* XVIII.
- (6) KLÜVER, H. (1933). *Behaviour Mechanism in Monkeys*. Chicago University Press.
- (7) KÖHLER, W. (1915). Optische Untersuchungen am Chimpansen und am Haushuhn. *Abh. preuss. Akad. Wiss.* III.
- (8) LOCKE, N. M. (1935). Colour constancy in the rhesus monkey and in man. *Arch. Psychol.* no. 193.
- (9) MARTIN, T. M. & PICKFORD, R. W. (1938-9). The effect of glare on apparent size relations. *Brit. J. Psychol.* XXIX.

(Manuscript received 30 September 1950)

CRITICAL NOTICE

By F. W. WARBURTON, *School of Education, University of Manchester*

The Structure of Human Abilities. By P. E. VERNON. London: Methuen. 1950.
Pp. viii + 160. 12s. 6d. net.

This book brings together all researches published during the last fifteen years which have employed factorial techniques to study abilities, and gives the reader some acquaintance with the principles of factor analysis and with the divergent views of British and American investigators in this field. This material is used to build up a new picture of mental structure.

We have long needed a critical survey of the applications of factor analysis in the educational and vocational fields. Prof. Vernon carries out this task with extreme thoroughness and competence. An examination is made of investigations into intellectual abilities, educational attainments, occupational aptitudes, mechanical skills, the aesthetic, psycho-motor and physical capacities, sensation, perception and imagery. This prompts the discussion of a large number of practical problems, such as the influence of selection procedures on teaching in primary schools, the effects of practice on test scores, speeded tests, factors in mental deterioration and the relation between cultural standards and performance. Some of the conclusions reached are contrary to earlier opinion. It is suggested, for example, that the tendency reported in previous investigations for intelligence to be differentiated into more specialized abilities during adolescence and early adulthood is illusory, the appearance being due mainly to greater selectivity in the older groups studied. It is also considered that the children most likely to benefit from advanced schooling are those with the best attainment, not those with the highest ability, as was formerly believed.

Throughout the book great stress is laid on the importance of keeping close to the experimental evidence, and the author strongly urges the statistical approach to the educational and vocational areas of research. Whether psychology is a science is perhaps a matter of definition, but he has no doubt that in our studies of ability the use of scientific methods has been justified. He holds that whether an ability depends on some power of the mind is of interest to metaphysicians but not to scientists, who are concerned with the measurable activities of human beings, and that it is better to conduct experiments than to theorize about psychological origins. In his view statistical considerations must often take first place in practice as well as in theory; he thinks, for instance, that it may be a waste of time and money to give tests of low validity, even if they are attractive to candidates or employers. The disadvantages of analytic techniques are pointed out, and attention is drawn to the limitations of the rôle that factors can play in exploring the personality.

The more theoretical part of the book also gives the first popular account of the hierarchical group-factor theory and contrasts it with Thurstone's work on the primary abilities. There is an increasing tendency for British psychologists to regard abilities as a system of hierarchical factors of various degrees of generality, the more comprehensive including the more specialized, as countries include counties. 'Genealogical' tables, headed by 'g', can now be drawn, attributing different levels to the various factors and to their 'descendants'. A structure of this kind has the merit of clarifying the relation between the hundred or so factors already reported by research workers and, in this respect, is a great advance on earlier theories.

Prof. Vernon suggests that it follows from the notion of hierarchical structure that a person's score in a battery of tests, of minor factors, such as '*v:ed*' and '*k:m*' may be summed to form an estimate of a major group factor ('g') without actually giving a 'g' test. It is difficult to see why he explicitly attributes this advantage to the group-factor theory, unless the hierarchical factors are correlated with one another. This is an important point. Nearly all British psychologists prefer to keep their factors uncorrelated, which means that adding a person's scores in certain factors will not indicate his score in others. The case is different for test scores (as distinct from factor scores), since tests are correlated and may be summed. The practice of summing test scores that is recommended in this book may thus afford some reconciliation with those American workers, who use correlated factors and, on occasion, sum them to give a composite measure of general intelligence.

The Thurstonian may protest that his standpoint is rather sketchily given. He will miss the case for oblique factors: for simple structure as the guiding principle to questions of the generality and orthogonality of factors; and for the centroid 'g' being an average, dependent on the particular tests that happen to have been included in the battery. American writers are even said to have adopted a 'neo-faculty' theory. This seems a little hard, whether one contrasts factors with faculties (Chapter I) or takes them as identical

chapter V). Yet the author's belief that it is both psychologically foolish and mathematically difficult to little 'g', as the Americans often do, seems the best hypothesis for practical purposes, and he is surely right in protesting against 'primary' factors of relatively low variance which depend largely on differences the background of the populations studied.

There are two footnotes of considerable interest to the statistician. One gives the opinion that the multiple-regression technique is too efficient, since it prohibits the inclusion of rather similar tests which might improve the reliability of predictions; and the other recommends formulae for testing factor significance, which compromise between British caution and American expansiveness.

The structure of the abilities is illustrated by some excellent diagrams, although the prominence given some of these (and in the body of the book) to Alexander's 'X' may suggest that our knowledge of personality factors is sounder than it actually is, and that these factors can be compared on equal statistical terms with those established in the cognitive field.

Criticism is apt to give a false impression of a book as comprehensive and lucid as this. It should appeal to everyone interested in psychology, and to the specialist it is indispensable. Prof. Vernon has made a major contribution to the literature of educational psychology.

PUBLICATIONS RECENTLY RECEIVED

L'Année Psychologique. Ed. H. PIÉRON, A. FESSARD and P. FRAISSE. 1948. Paris: Presses Universitaires de France. 1950. Pp. viii + 700. 1800 francs.

There is a story to the effect that Prof. Piéron was once asked why his book *La Sensation: Guide de Vi* contained no bibliography, and he replied: 'How could I put one in?—It would simply consist of several volumes of *L'Année Psychologique*!' While one might feel some surprise that even several such bulky tomes could contain a catalogue of Prof. Piéron's enormous erudition in the field of sensation, his remark scarcely does justice to the great and varied interest in psychology as a whole which is yearly brought to a focus in *L'Année*. The Paris school still derives its core of inspiration from the experimental traditions—healthy astringent ones—laid down by Binet and his associates, and later so richly developed by Piéron himself. One may gain some idea of its present wide scope and intensive industry from this issue of *L'Année*. And one has once again the opportunity to appreciate that peculiar and happy blend of enthusiasm and discrimination which only the French can commonly command.

As usual, the contributions fall into three classes. There are original papers, topical reviews and analytic bibliographies. This diversity of content and intention is favourable to the reader's alertness and continued interest. Too often annual reviews dismally batter, and finally drown, him in a featureless succession of waves of murky erudition. Among the original papers are to be found valuable, if perhaps not always arresting, communications on a variety of subjects, from Galifret and Piéron on intra-retinal transmission, to Bernier on aptitude tests. Popov has some interesting electro-encephalographic observations relevant to cortical processes underlying the establishment of conditioned reflexes, while Reuchlin describes a formal scheme of notation for the recording of behaviour during performance tests, such as the Koh's Blocks. Comparative psychology is represented by the contributions of Viaud, and of Biatte, and the more classical topics of perceptual rhythms, of attitudes in learning, and of recognition are among others not neglected.

Among the topical reviews are several which will strike a number of readers as very timely and useful. Such are those of Livingston on the functions of the frontal lobes—with an extensive bibliography—and of de Montpellier, who provides a freshly critical discussion of the conditioned reflex in its relation to learning. Prof. Piéron himself surveys recent work on pain (a subject all too often neglected by the psychologist) in which neuro-surgeons have recently thrown up a quantity of interesting material. He also gives us an extremely useful summary of recent developments in the establishment of scales of sensory magnitudes. Reuchlin considers closely the functions of the psychotechnician—a problem which evidently demands in France just as much attention as it ought to be given in this country.

The analytic bibliographies do not of course replace *Psychological Abstracts*. They are selective and intended to be read rather than referred to. They succeed in providing pointers in a variety of fields.

It is clear that present-day French psychology is second to none in thoughtfulness and penetration. This book is the work of men who write from the love of their subject rather than with the aim of advancing its claims at large. The breadth of interest, which might have seemed in temporary eclipse fifteen years back, has re-established itself. It may be that the combination of qualities shown in *L'Année Psychologique* represents a real condition for the future of the subject—possibly a vital one.

R. C. OLDFIELD

The Psychology of Sartre. By P. J. R. DEMPSEY. Cork University Press (Oxford: Blackwell). 1950. Pp. xii + 174. 12s. 6d. net.

Dr Dempsey devotes two-thirds of his book to giving an exposition of Sartre's psychological theories. In the remainder, Part II, he offers brief broad criticisms and 'principles of solution' of the problems involved.

Psychologists and those generally interested in Sartre should be glad of Part I. After opening with an account of Sartre's sources, it provides a painstaking and thoroughly annotated account of his theories, adhering closely to Sartre's own idiom. Thus, within the compass of about a hundred pages, Sartre becomes more readily accessible to English-speaking readers; at the same time there is a certain amount of explanation that should help us a little towards understanding Sartre. There is no doubt that the author has steeped himself in his subject and is at home with it.

This is not to say that Sartre is now made clear. Although Dr Dempsey has expounded Sartre more clearly than Sartre himself has done, and although it is useful to have such an exposition in Sartre's own idiom, Dr Dempsey has unfortunately refrained from translating the resultant product into an intelligible

language of some standard kind. Hence the reader will still have to put in a great deal of laborious work to find out what Sartre wishes to say. The exposition calls out for metalinguistic clarification (not as a substitute for what is given but as an addition).

This would expose Sartre if what he has to say is essentially trivial; it would be a signal contribution to psychology if—as is probable—Sartre has some important things to say.

Dr Dempsey's criticism of Sartre's doctrines consists of rejecting them rather than of sifting them. Perhaps it is in consequence of this that he has overlooked the most important part of Sartre's theory of emotion; Dr Dempsey mentions that emotion is conceived as magical, but omits to mention Sartre's account of the mechanism of this, to do with the movement of images.

Dr Dempsey bases his brief account of some 'principles of solution' (by which he appears to mean guiding ideas) upon Aristotle and Aquinas; but, unless the problems are first made precise, solutions or even principles of solution would seem to be premature.

One jarring note must be alluded to. Dr Dempsey cites a writer on depth psychology as saying that Sartre's own emotional life was disordered and that 'his emotional principles spring from these interior deviations' (p. 24). The next sentence, which appears to be Dr Dempsey's own, describes Sartre as 'a neurotic genius . . . of brilliant intellectual and of immature affectivity . . .' Dr Dempsey has not, however, convincingly shown the relevance of these remarks. Moreover, the evidence given in support of them is quite inadequate (it consists of a reference, brief biography (pp. 22-4), and a page of assertions (p. 148) about Sartre's character). In these circumstances, derogatory remarks are out of place. J. O. WISDOM

Masks of Love and Life. By HANNS SACHS. Edited with an Introduction, Memoir and Glossary by A. A. ROBACK. Cambridge, Mass.: Sci.-Art. Publishers. 1948. Pp. 303. \$4.25 net.

To appraise this book briefly but fairly is an exceedingly difficult task, and that is one of the main reasons for this very belated review. Dr Hanns Sachs was by original profession a lawyer and had no medical degree. Nevertheless, he became a member of the staff of the Harvard Medical School, and achieved a wide influence in America, being, it seems clear, much loved by those who knew him best. In any case, it is hardly ever possible to judge a doctor fairly in terms of what he may write alone, for his real power, if he has any, is apt to find its chief expression in his personal contacts. Dr Roback, in his brief but penetrating introduction, presents Sachs as a 'wise' man. Wisdom is perhaps not very likely to spill over into anything very exciting and memorable in words. For that, very often at any rate, demands the kind of oneness which wisdom does not have. So the wise man's reputation is found more often in practice than in theory.

Sachs, so Dr Roback says, threw his weight on the side of working for a recognition of the importance of psycho-analysis in non-medical fields. For him, more than for most 'psycho-analysts, psychology meant, above all, the means to inquire into the daily behaviour of human beings'.

In so far as the essays printed posthumously in this volume are concerned, Roback is unquestionably right. They are admirable essays, contemplative and not experimental, or even, in the usual sense of the word, clinical. They are gentle, pleasant and wise. If they see men with masks of love and life, it is with eyes that are tolerant and hands that have no raging itch to strip the masks away. The titles are not very informative: *Peering over the Fence*; *Locked in a Room with Open Doors*; *The Mastery of Shadows*; *The Variables*; *The Path to the Tree of Life*; sixteen in all, of greatly differing lengths. It is also far from easy to say, in sharply defined language, what is their message. But it is likely that anybody who reads them without prejudice will be helped to achieve some comfort and tolerance, and a sort of understanding that cannot be readily expressed, of the affairs of mankind.

The Albert Schweitzer Jubilee Book. Edited by A. A. ROBACK. Cambridge, Mass.: Sci.-Art. Publishers. Pp. 508. No date of publication or price given.

The book is a presentation volume of Albert Schweitzer on his seventieth anniversary. It is a prolonged piece of hero-worship, of varying interest and technicality. Its sections are entitled: *Personal Interpretations*; *Methodology and Evaluation*; *Aesthetics*; *Ethics and Civilization*; *Humanities*; *Medicine*; *Theology and Religion*; *Epilogue*. The whole forms a fine critical appreciation of Schweitzer and his work, but it will be seen that, as was indeed inevitable, the bulk of it lies outside direct psychological interests.

Compulsion and Doubt. By WILHELM STEKEL (2 vols.). London and New York: Peter Nevill. 1950. Pp. 648. 36s. net.

The unfortunate death of Dr Stekel in this country, to which he had escaped from Austria, robbed us of a very charming, not to say naïve, personality and an original clinician. For though his theories largely followed those of Freud, his approach to these problems of the neuroses was fresh and his method of treatment original. For one thing, his diagnosis of the patient's illness, through dream interpretation, depends more on intuition than on the strict interpretation of dream symbols.

Stekel had a saying with regard to himself in relation to Freud, that the dwarf who sits on the shoulder of the giant sees further than the giant—'not', he would add, 'that I consider myself a dwarf'. These books carry out that theme, for though they do not contain the acumen of Freud's theory, they describe many examples of compulsions with a freshness and originality which make easy and pleasant, as well as profitable, reading.

As he is so close a follower of Freud, it is somewhat surprising to find that he does not believe in the unconscious. 'I myself, though standing on the ground of psycho-analysis, do not believe that there is such a thing as a psychic unconscious' (p. 228). He considers that those patients suffering from compulsions know the causes of their trouble but blind themselves to them, and need to be brought to face them. Memories of these unpalatable experiences are only temporarily annulled, but the patient is at all times capable of remembering them. A girl, for instance, who has phobias of any sharp object for fear she is deflorated and of losing her virginity, was not really guarding herself against the loss of her virginity, but guarding herself against the memory of a seduction, for she was bluffing herself and others that she was a virgin when all the time she had allowed herself to be seduced, a fact elicited by Stekel from a dream and admitted by her when directly confronted with the fact. In this respect Stekel's practice seems to correspond to that of Jung, who regards the neurosis as due to present-day moral conflicts, although Jung, of course, has a very firm belief in the reality of the unconscious.

The subjects dealt with in these books are: The Unconscious in Compulsive Neurosis; Doubt Homosexuality and Patricide; Obsessive Neuroses; Schizophrenia and Compulsion; Criminal Impulses. J.A.H.

The Foundations of Social Anthropology. By S. F. NADEL. London: Cohen and West. 1951. Pp. xii + 426. 25s. net.

This book is frankly about method in social anthropology. It is concerned with the 'logical premises that underlie our knowledge of societies (whether they be primitive or otherwise) and with the pre-requisites conceptual and technical, of any enquiry meant to lead to this knowledge'. It seeks to make explicit and to evaluate the considerable body of axioms in method and empirical 'co-variations' which, while generally relied on and already productive of results, has to a considerable extent been tacitly and unguardedly developed in the course of anthropological field studies. Prof. Nadel does not underrate his task, which he sees as no less than to organize anthropological knowledge 'in a new order [and] in the form of a system'.

The data of social anthropology are approached with reference to the structure and function of institutions, those complexes of integrated behaviour patterns and norms in which cultural behaviour and social interaction are canalized, and to description of which so much attention has been given by sociologists and, in Britain particularly, by social anthropologists. The conceptual and other problems arising in an analysis of institutions are, however, considered with reference to the wider framework of our understanding of action in society. He brings the generally accepted canons of scientific method to bear on the procedures and assumptions involved in establishing, through field observation, recording and systematization, the regularities of behaviour postulated for any given 'Society'. Stressing the pervasiveness of signs and symbols, including ideas and their symbolic representation in the adaptation of behaviour in society, he shows that these must be regarded as integral to and inseparable from social activity. His extended analysis of the relations between 'institutions' and 'groups' is carried out with reference to two aspects or 'orders of things social', derived respectively from a consideration of the 'aim contents', including the purposive character of action patterns, on the one hand, and, on the other, of the relations between 'persons', i.e. social positions that may be occupied by individuals. Showing that these two aspects underlie the much, and often loosely, discussed dichotomy between 'culture' and 'society', he not only stresses the substantial interdependence between these two orders, and the need to recognize that every 'social fact' is relevant at once to both of these conceptual fields, but also that the significance of any social fact in the two orders may vary with some degree of independence not only as between 'Societies', but even within a single 'Society'. He also agrees with those who see the need, sometimes obscured by linguistic difficulties,

to distinguish between the levels of abstraction involved in these respective projections of total group behaviour on to the dimension of action (culture) and on to the dimension of relations and groupings (society or social structure). He recognizes that the latter, which generalizes beyond the qualitative character of the patterns of behaviour, the 'aim contents', etc., systematizes at a further level of abstraction concerning the relative position of categories of factors and their co-ordination. This, it may be suggested, is worth close consideration by psychologists interested in the establishment or transmission of value systems, since failure to perceive the nature of the distinction between action patterns and systems of social relations and the introduction of unwarranted inferences from the one to the other would appear to be a defect to which some generalizations in the field of 'personality and culture' are prone. Elucidating the distinction and relation between the concrete individual and the 'person' as a social role and position open to a number of individuals, Nadel implicitly warns against the assumption that, in differentiated societies there are, or can be, generally operative social processes leading to a comprehensive norm of social personality'. Since there can be no single predominant combination of roles warranting recognition as a single or pervasive social personality, '... we should be referring to a synthetic personality to which nothing corresponds in practice and to a hypothetical (and internally inconsistent) individual whom we made into a summary of all the social persons existing in the society... In defining the [varieties of] social personalities we should in fact be describing the whole working of a culture or society...'

This book should have a wide interest for psychologists concerned with the social context of mental functions. Although it is not a study of the relations of the two disciplines as such, it takes up a number of further questions concerning psychological processes involved in social activity. Its importance for the psychologist lies more in the fact that this attempt to systematize the categories developed by social anthropology for the study of structure and function in small-scale societies, explores fundamental problems concerning the nature of our knowledge of social process, and makes explicit the dependence on psychological principles of many sociological formulations and explanations. That postulates concerning basic psychological processes and the variables affecting them underlie any interpretations of social activity may, perhaps, be obvious, but rigorous examination of this nexus is often lacking and much 'social psychology' suffers from lack of fit psychological and sociological categories. Prof. Nadel's study should, therefore, apart from its cogent expositions of advances in social analysis, stimulate reconsideration of the inter-relations between the two fields.

DARYLL FORDE

Public Opinion in Soviet Russia. By ALEX INKELES. Harvard University Press (London: Geoffrey Cumberlege). Pp. xx+379. 32s. 6d. net.

Public Opinion in Soviet Russia describes 'the principal media of mass communication in the Soviet Union', and discusses the how and why of their use. This, Dr Inkeles shows, is dominated by a political philosophy which holds that social salvation can be attained only under the leadership of the Communist Party; people must therefore be got to believe and to do what the party leaders know to be best for them. This end may be sought through the media of mass communication, but not only this; the media are two-way channels that convey orders and instruction from leaders to people, and also information about the people to the leaders. The discussion of this second function, one intended to enable the leaders to keep in touch with, and to control the pace and execution of their policy, contains much that is both new and interesting.

Of the four media dealt with by Dr Inkeles, Oral Agitation, the Press, Broadcasting, and the Film, the first two have this double function. The Agitator, a sort of part-time and low-level personal relations and liaison officer, wins support for the leaders' policies through his personal contacts with individuals and groups, and keeps the leaders posted about the attitude of his charges. His position is far from easy. Not seldom he has to 'sell' heavier burdens and shorter commons to his workmates; and there are always the risks attendant upon bringing bad tidings to those in authority. Not surprisingly, the turn-over of personnel is large, and methods and personnel are now being tried that do not require such close contact. The discussion of the advantages and disadvantages of personal contact in influencing people is one of the most interesting parts of the book.

The penalties for statements construed as hostile—the very real responsibility that haunts the exercise of Soviet freedom of speech—must be still more obvious to those who are called upon to criticize the carrying out of the leaders' policy. There are the workers' and peasants' correspondents, who are not professional journalists but party members or sympathizers, who from time to time write letters to the local or national press, criticizing the state of affairs in their locality; and above all there are the editors. These have to venture oftener and more widely. And if there is danger in too much, one may be purged for doing too little.

A first-hand psychological study of these agitators and editors is as tempting as it is impracticable. But it is only one of the many fascinating psychological issues indicated by Dr Inkeles in a work that is not only indispensable to the student of public opinion but also of great general interest. He makes clear how and why the Soviet leaders try to control public opinion, though for lack of data he cannot say how far they succeed. Dr Inkeles has succeeded, as far as it is possible to succeed, in a field bristling with man-made as well as natural obstacles.

H. E. O. JAMES

Soviet Psychiatry. By JOSEPH WORTIS. Baltimore, U.S.A.: Williams and Wilkins (London: Ballière, Tindall and Cox). Pp. xvi+314. 38s. 6d. net.

Most Soviet text-books begin with a political introduction, stating the ideological basis of their subject-matter, and quoting Lenin and Stalin in every other sentence. But after that they have enough sense to leave Lenin and Stalin and come to reality. To the psychiatrist, reality is his patient, and Russian psychiatric books, leaving aside unconscious dynamics, give excellent and vivid clinical pictures. Admirable symptomatology has been traditional in Russia, and may be partly due to the nature of the Russian mental patients who display their symptoms more freely than the patients in the West.

Prof. Wortis carries ideology throughout the book. Psychiatric conditions, he argues, reflect social conditions. Since he does not know these conditions from personal experience, he assumes them to be as reflected in the Soviet Constitution and similar declarations. Thus, 'the framework of child care is embodied in the Soviet Constitution and system of law on children and the family'; 'the family is a cherished institution in Soviet life'. Soviet education is an integral part of the planning for the welfare of the population 'which characterizes every phase of Soviet life'. The incidence of mental disturbance in the U.S.S.R. is assumed to be low (this is proved by statistics; the Moscow area has no need for more than 0.7 psychiatric hospital beds per 1000 population, while New York State has 7 beds per 1000 population), and the reasoning is simple: 'Socialist construction creates healthy conditions of life and work and thus eliminates the tensions which cause so many neuropsychiatric disturbances.'

Having depicted in great detail and scholarly manner the Soviet achievements in psychiatry (many of which, especially in neurophysiology, are certainly valuable), Prof. Wortis quotes a Soviet psychiatrist's suggestion to liquidate capitalism and create the socialist society: 'Therein lies the promise of the future health of man.'

The preparation of the book was supported by the Harvard Russian Research Center. T. GRYGIER

Bilingualism. An inaugural lecture. By PAUL CHRISTOPHERSEN. London: Methuen. 1949. Pp. 16. No price given.

Prof. Christophersen's inaugural lecture delivered at the University College of Ibadan (Nigeria) on Bilingualism dealt in a masterly and provocative fashion with the psychological effects of bilingualism and with the main problems of language study in bilingual countries. His findings, which are suggested generally by illustrative examples rather than compelling but boring statistics, should be carefully studied by education authorities.

Professor Christophersen's main theses are:

(1) It is not possible for an ordinary human mind to handle more than two—or at most three—languages with any degree of perfection, to express more than the same rather limited set of notions in many tongues.

(2) There should be a clear understanding of the term bilingualism. 'A bilingual is a person who knows two languages with approximately the same degree of perfection as unilingual speakers of those languages.' According to Christophersen, a bilingual is 'never 100 % efficient in either language'; and of course he may be more at home in one language than the other.

(3) A bilingual person's two languages need not have been acquired at the same time. After early youth, however, most people lose the ability to learn a foreign language perfectly, but some exceptional people retain a linguistic 'virginity', a power to retire into the inarticulate and emerge again in the foreign language they are learning'.

Prof. Christophersen in his address is frequently concerned with the 'linguistic efficiency' of bilingual speakers. This new statistic is not defined. What factors are involved in it and how are they measured? These and other questions remain unanswered. Accent, fluency, phrasing, idiom and vocabulary are all considered by Prof. Christophersen. Is each involved in linguistic efficiency? Is perfection or 100 % efficiency a universal objective standard, or does it refer to an individual's estimated potential mastery of the linguistic factors mentioned? The latter and vaguer connotation would appear usually to be the more

appropriate interpretation of its use by Christophersen. But 'linguistic efficiency' may yet develop into a new statistic based on factors that are all objectively measured. Objective vocabulary tests exist; objective tests of the other factors are not inconceivable.

Prof. Christophersen's findings on the psychological effects of bilingualism are interesting.

(1) Bilingualism tends to reduce the capacity for verbal memorizing since experience tends to be retained as a sensory image, leaving the expression of that experience to whatever linguistic form is called for at the time of telling.

(2) The absence of an immediate verbal link between experience and language leads in some cases to hesitancy in speech.

(3) The bilingual's two languages are far more than a medium for the expression of ideas. They represent two worlds, two modes of thought, two codes of behaviour, two emotional attitudes to things. Unless the bilingual has *deliberately* analysed the two ways of life, Christophersen considers, with Lagorde-Quorst, that bilingualism brings the danger of 'split personality', makes the bilingual person unpredictable to himself and to others. A shift of sympathy from one group to the other, as well as environmental change, leads to a change in the preferred language.

(4) A bilingual does not normally translate from one language to another. The two languages are kept apart in the mind. The art of translation has to be learnt specially.

(5) The automatic bases of speech may achieve perfection; and accent, governed by reflexes, is least likely to be forgotten or mixed up.

(6) A bilingual is critical of verbal forms, less likely than unilinguals to attribute meaning where there is none.

(7) Bilingual children are usually retarded. The mental effort required to learn a second language deflects mental energy from the acquisition of other skills.

Of these effects, only the third seems overstated. One consequence should surely be that a higher percentage of schizophrenic abnormals would be found among bilinguals than unilinguals, a suggestion which the reviewer has been unable to confirm.

Unpredictability from a unilingual's viewpoint may be any questioning of the extent of individual loyalty to the unilingual's group. Such questioning may arise from a strong sense of individual rights and duties as well as from a sense of conflicting group sympathies. This strong sense of individual self-sufficiency among some bilinguals the solution to the dilemma of conflicting loyalties. Just as he turns from one language as the answer to the linguistic dilemma, so he turns from both as the answer to rival group claims on his loyalties. This solution is not mentioned, however, by Christophersen.

Prof. Christophersen only takes up incidentally the question of whether the bilingual child's retardation is corrected later in life. Malherbe's claim, that it is, is not mentioned; but by considering the achievements of gifted, adult bilinguals he comes to the conclusion that bilinguals have probably got more than their share of great writers. The tendency of bilinguals to gravitate towards science and art is explained as a form 'of escapism from word bondage'; and it is implied that in these fields bilingualism may be advantageous. Christophersen remarks that 'the extreme sensitiveness and artistic temperament of the Celt' may be attributed to bilingualism. These considerations are all in answer to the question 'Is bilingualism desirable?'

Shrewd comments are offered on the effects of schooling upon the language of children. Christophersen finds that, although parents may insist upon speaking their own language in the home, the children, in most cases, will come to feel more at home in the language of the school, a language that was originally a foreign tongue to them.

This comment is significant. It reiterates a truth well attested in Wales where, since 1848, and up to recent times, little or no attention was given in schools to the mother tongue. Many educational authorities in Wales now give the first instruction in Welsh in those areas where that language predominates, and even in English-speaking areas attempts are being made to provide Welsh schools for those whose home language is Welsh.

Prof. Christophersen attacks old methods of language teaching based on links being forged between words in the mother tongue and in the foreign language. The danger is a second language manipulated according to the rules of the mother tongue. Indeed, many immigrants illustrate the danger in their broken language, which persists, although the mother tongue is forgotten.

Turning to national language policies in the latter part of his talk, Prof. Christophersen first emphasizes his belief 'in the essential sameness of the human mind everywhere' and pleads for 'a little less disunity in the world to-day'. English can be used, alongside the vernacular languages, to express the spirit of a country. He does not look with favour upon attempts 'to create linguistic individuality where previously there was none'. He warns especially of the dangers of abandoning the vernacular languages too quickly

and completely in favour of the second language; the resulting language may be a national parallel with the peculiar language of some immigrants. National extinction too is possible. And, thirdly, the decision to change a nation's language is so momentous, linked as it is with national tradition, that no one generation should take the responsibility.

These fears need not afflict Welshmen, for the Welsh Department of the Ministry of Education is well aware of the claims of the Welsh language, as are local educational authorities in Wales.

DAVID J. STEPHENS

The Scottish Council for Research in Education. XXXIV. *Studies in Reading*, vol. II. University of London Press. 1950. Pp. xvi+113. 7s. 6d. net.

This publication consists of three reports of researches undertaken in partial fulfilment of the requirements for degrees in Education and an interesting analysis of disabilities in reading and their remedial treatment by the late Gregor Macgregor, Director of Education to the Fife Education Committee and a foundation member of the Scottish Council for Research in Education.

The first study, 'Socio-Economic Status and Reading Ability—A Study in Infant Reading', by Violet M. McLaren, attempted to determine the relationship between the socio-economic level of the home and progress in reading. Another phase of the inquiry assessed the efficiency of two methods of teaching reading which are not adequately described. In one method, 'the initial approach to reading is made through stories. Parallel with this activity, although wholly distinct from it, the pupil is given systematic practice in phonic word recognition which he later applies to phonic words introduced in his reading. Short sentences are memorized and then the individual words. The second method is 'also a combination of phonic and sentence methods but there is less emphasis on preparation for reading. Phonic training and reading for meaning proceed simultaneously. A few letters are learned and short words are formed, the initial blend being used as in the former method.' The results indicate no significant differences in the efficiency of the two methods. In agreement with previous studies, a significant relationship between socio-economic status and progress in reading was indicated and, in this study, was consistently maintained from $5\frac{1}{2}$ to 8 years.

The second study, by Christian D. Taylor, 'The Effect of Training on Reading Readiness', applied to 114 children of mean chronological age 6 years and 3 months and of varied social background in Fife and Edinburgh, The Metropolitan Reading Readiness Tests for Kindergarten and Grade 1 and The Metropolitan Achievement Tests Primary 1 Battery, Form A. These tests were standardized for American children and in this study it was assumed that the average American child of 6 years and 3 months would record a reading age of 6 years and 3 months. The average I.Q. of the Scottish children was 100.59, $\sigma = 13.68$ (Terman Merrill), and since they had entered school a year previously, they provided a basis for comparison in respect of skill in reading with American children, who are claimed to enter school at 6 years of age. The mean score of the Scottish group on the reading readiness test was significantly higher than the American norm and, on the achievement test, they recorded a reading age of 7 years 5 months, again significantly above the American norm. The author does not regard the results as conclusive, but submits that reading readiness is not so dependent upon maturation as is customarily assumed.

The third paper, 'Children's Leisure Reading Interests', by Doris C. Dunlop, reports evidence of definite choice by children, of the influence of the radio in determining popularity and a preference for fiction. There are some interesting speculations on the retardation of the development of the child's literary taste by the shortage of children's classics during the years of war and the aftermath.

All three papers give a very useful survey of the literature associated with each problem.

F.V.S.

Recent Advances in Diagnostic Psychological Testing: A Critical Summary. Edited by MOLLY R. HARROWER. Springfield, U.S.A.: Charles C. Thomas (Oxford: Blackwell Scientific Publications). 1950. Pp. x+120. 20s. net.

This book on diagnostic testing is written by six separate authors. The first contribution is by James G. Miller and deals with contemporary trends in psycho-diagnosis. In this he discusses, in a valuable way, changes in outlook which have taken place in recent years. The second contribution is by Robert E. Harri and is entitled 'Psychodiagnostic Testing in Psychiatry and Psychosomatic Medicine'. The third contribution is on neuropsychiatry by Hans-Lukas Teuber. The fourth is on the application of diagnostic psychological methods to counseling and psychotherapy by George A. Meunch. The fifth is on recent development in diagnostic testing of children by Lawrence J. Stone. The sixth and last contribution is on test construction

and methodology by Joseph Zubin. The reader will find this book a useful review of present-day tests, points of view and methods, and it will serve to bring his ideas up to date on the problems of diagnostic testing with which it deals. There is no index, but there is a useful bibliography attached to each chapter.

R. W. PICKFORD

Guidance of American Youth. By JOHN W. M. ROTHNEY and BERT A. ROENS. Harvard University Press (London: Geoffrey Cumberlege). 1950. Pp. xvi+269. 32s. 6d. net.

General Clinical Counseling. By MILTON E. HAHN and MALCOLM S. MACLEAN. New York: McGraw-Hill Book Co. Inc. 1950. Pp. xii+375. 30s. net.

The practice of 'counseling' or 'guiding' school children and students towards the work which appeared to be most suitable for them originated in the United States some time before the war. It was found particularly valuable in view of the very widespread choice of educational courses offered, the frequent change of subjects studied, and the general absence of any standardized curricula like those of the English grammar school. Thus American adolescent and adult students were often largely ignorant as to where their capabilities lay, what subjects they could study, and what career they should aim at. The work of the 'counselor' was to ascertain what the student wanted to do, what he could do, and how best he could do it. This work received a further stimulus after the war, and expanded considerably, in advising and assisting ex-Servicemen on how to obtain suitable forms of educational and occupational training.

The procedures adopted appear still to vary considerably from place to place, and according to the age of the individual guided. They are not dissimilar from vocational guidance procedures in this country; but they are generally performed by a staff of 'counselors' working in a University, college or school, or a group of these. The American procedures involve a more detailed study of the individual, in tests, records of grades, personality assessments, and numerous successive interviews. Indeed, the documentary material accumulated about each individual must be formidable in amount. It does not appear that the understanding which is gained of the individual and his problems is much greater than that obtained by the English procedure. But much time and effort is spent, and probably well spent, in giving the individual the opportunity to talk out his troubles, and in advising him how to tackle his particular problems and adjust himself to difficulties which he might otherwise find insurmountable.

Of these two somewhat expensive books, *Guidance of American Youth* is devoted mainly to a small-scale intensive study of the procedure and effects of 'guidance' in schools. It gives the reader a clear picture of how this was carried out, and how it led to the general adoption of a similar 'guidance' procedure throughout a group of schools, with children of all ages. *General Clinical Counseling* gives a wider view of 'counseling' at the University stage—particularly necessary in view of the fact that there may be more than 2700 different courses of study available at a single University. But the specimen procedure described appears excessively elaborate and time-consuming, and also distinctly vague and lacking in definite conclusions.

The methods adopted in 'counseling' may seem to us at first sight somewhat repellent in the amount of interference they may cause, both in the individual's personal affairs, and also in the teacher's planning of his work and his relation with his pupils. We hear, for instance, that 'practically all pupils are exposed to classes in occupational information in the eighth and ninth grades'. But this reaction may be a mere prejudice surviving from the days when it was thought that parent and teacher between them could guide the child effectively. Perhaps a good careers master and a good tutor *could* do all that is done by the 'counselor'. But do they?

M.D.V.

Le Test du Labyrinthe. By F. CHAPUIS. Berne: Hans Huber. 1949. Pp. 142. Fr. 12.50 net.

The author has devised three mazes similar to the adult mazes of Porteus, though two at least are more difficult. The test is administered twice—the first time with emphasis on accuracy, the second with the emphasis on speed. The change in set makes little difference to the performance.

The test is rightly considered to assess certain aspects of intelligence and character. It seems unfortunate that a global score is finally derived which obscures much of the psychological meaning of the performance.

There are many shrewd observations of the manner in which the test is performed by different individuals. Some of these observations might profitably have been quantified.

The practical value of the test for personnel selection for the Swiss Federal Railways is nevertheless well established.

G.A.F.

Work Adjustment in Relation to Family Background. By JEANNETTE G. FRIEND and ERNEST A. HAGGARD. Stamford University Press (London: Geoffrey Cumberlege). 1948. Pp. xvi + 150. 14s. net.

This is a detailed report on eighty selected cases passing through the Vocational Counseling Service of the Family Society of Greater Boston between 1934 and 1943. All had at least six, some fifteen talks with the counselor over periods of from four to nine years, and often the records of their parents, sometimes of their grandparents, were available for reference. The district case worker supplemented the interview by talks with them and their families.

Impetus was given to the research by the flooding of the service with applicants during the depression years and the need for selecting those most likely to profit by it. Placement according to needs, aptitudes, skills and interests had only too often failed because of the overriding influence of factors related to the individual's underlying emotional attitude to his work and the unconscious demands he made of any given job situation. The family background might hold the clue to these elusive factors, but little was known about what specific items in early life might be important in determining personality, work reactions and job adjustment, and how these interacted with given factors in the job environment.

The investigators used a rating schedule of 173 items, grouped according to Early Life, Mature or Family Life, Early or Beginning Jobs, Response to Counseling, Personality Patterns and General Work Reactions, Reactions to Specific Working Conditions, General Work Capacities, Adjustment and Improvement. Two trained persons, consulting the case worker over difficulties, took an average of 130 hours rating each case.

Detailed statistical analysis of the results brought out some not always unexpected but interesting points, e.g. the importance of emotional security. Those rated 'high' for job adjustment came from families that were 'closely knit and strongly unified', contrasting strongly with the 'disorganization and lack of togetherness' in those rated 'low'. Similarly, antagonism towards parents was characteristic of the 'lows', less of the 'highs'. In authority situations the individual's attitude towards his supervisor closely resembled that which he had towards his parents while he was growing up. Favourable school experiences of any kind, scholastic, social or athletic, were more frequent amongst the 'highs'.

The report concludes that a rating schedule filled in by a team of competent raters is a useful diagnostic tool coupled with detailed family case histories, and that vocational counselors should give careful consideration to the underlying emotional needs and drives of the applicant and be better versed in the work conditions likely to satisfy them, as the psychological climate of a job is of paramount importance.

MARGARET GREGORY

Manpower Mobilities. Bulletin 10, October 1950. Industrial Relations Centre, Minneapolis. University of Minnesota Press (London: Geoffrey Cumberlege). Pp. iv + 56. 75 cents.

The two short reports included in this bulletin are likely to be of interest chiefly to fellow-workers in the field of investigation of manpower mobilities. The aim in each case was to make an intensive study of a local urban labour market (one in St Paul in 1941-2, the other in Minneapolis in 1948) with a view to delineating groups and establishing trends and trying out research methods for future more widespread surveys. The basic assumption, so different from that underlying vocational counseling research, is that modern industry requires a leaven of 'mobile' participants, and that it is necessary to determine if possible what in fact are the characteristics of this desirable element in the labour market. The results in this respect are inconclusive. The chief value of the work is in the progress made in the definition of terms and categories and the demonstration of the inadequacies of the methods employed.

M.G.

Modern Abnormal Psychology, A Symposium. Edited by W. H. MIKESELL. New York: Philosophical Library. 1950. Pp. 880. \$10.00.

This work comprises twenty-four essays by psychologists and psychiatrists on various aspects of abnormal psychology, viewed usually though not invariably from a Freudian standpoint. The dust-cover says that the book is designed to meet the interest of the general public in abnormal psychology, as well as the requirements of the student for a reference and source book. It would appear, however, to be neither sufficiently simple, clear and consistent for the former purpose nor scholarly and complete enough for the

ter. The symposium method has distinct dangers for relatively elementary as opposed to truly encyclopaedic works and they have not been avoided here. Besides overlapping between chapters and rather disturbing shifts of emphasis, there are complete contradictions of statement by different writers. For example, in one chapter the editor himself tells us: 'There is little evidence for the Id and there is not as much evidence for a conflict between the unconscious and the conscious as their exponents suppose.' This chapter is sandwiched between others in which the full dogma of the unconscious is pronounced *ex cathedra*. Such a discrepancy must surely trouble the attentive reader, but it is left unresolved. Despite claim to be comprehensive and exhaustive, the work ignores or cursorily skims over whole fields of current importance to abnormal psychology, such as the effects of brain injury and organic brain disease, the convulsive and surgical treatments, the genetics and biochemistry of mental illness, electrophysiology and the E.E.G., disorders of language and thinking, and the social or cultural factors involved in maladjustment. The bizzareries of 'psychosomatic medicine', however, are fully presented with absolute neglect of the mounting evidence against them. Proof throughout most of the work is by authority. Experimental or statistical data are rarely quoted. Thus Dr M. R. Sapirstein writes on the 'Therapeutic value of Psycho-analysis' and adduces not one single consideration to convince a sceptic that it has such value at all. One or two contributions stand out from the rest as genuinely scientific in approach, particularly Dr R. R. Brown's excellent account of pharmacological psychology. The clarity of several chapters is impaired by slipshod writing; Dr H. M. Graumann's study of perceptual and imaginal disorders seems to be a too-literal translation from German. Neither the general public nor students of any sort are likely to acquire from this symposium a firm understanding of abnormal psychology as a scientific discipline in rapid development.

J.P.S.R.

Culture in Crisis: a Study of the Hopi Indians. By LAURA THOMPSON, with a Foreword by JOHN COLLIER and a Chapter from the writings of BENJAMIN LEE WHORF. New York: Harper Brothers. 1950. Pp. xxvi+221. \$4.00 net.

For nothing else, this book would be valuable for the sake of some papers by Benjamin Lee Whorf, reprinted here in abbreviated form, on the structure of the Hopi language and its influence on Hopi thought. One could hardly have a better demonstration of the dependence upon language of our fundamental concepts and outlook upon the world. In particular, Whorf's account of the complete difference between our own and Hopi thinking about time is fascinating; as is his discussion of the treatment of cyclic and repetitive events in the Hopi language and its freedom from the spatial, quantitative metaphors on which we rely to indicate the duration, tendency and intensity of an experience. This part of the book is relevant to several aspects of psychology (besides physics and philosophy). It is highly stimulating, expressing subtle ideas in simple language.

The rest of the book gives a useful general picture of Hopi culture and the problems that arise in its adjustment to a dominant white culture. The book is one of an extensive series reporting an investigation of the social life of the Pueblo Indians of the south-western United States. Mistaken colonial administration and Christian missionizing of these peoples brought some of them (including the Hopi) close to disaster. A partial change of administrative policy, beginning in the early 1930's, led amongst other things to the establishment of the Indian Personality and Administration Research. This organization co-ordinated inquiries by workers in several of the social sciences and interpreted the results of their work in practical administrative advice (a summary of which in relation to the Hopi is included here).

Dr Thompson traces the history of the Hopi and surveys their economic and political position, giving a good account of the needless difficulties created by misunderstanding and intolerance on the part of administrators, medical authorities and missionaries. She shows the pivoting of Hopi culture between the 'female-centred kinship system' and the 'male-centred ceremonial system', which gave in its undisturbed state what is claimed to have been a satisfactory balance in the prestige of the sexes. In its account of the personality favoured by Hopi culture the book draws on the results of observations (including some testing) already reported in books and papers and much of it by now familiar. It is, nevertheless, very useful to have this material brought together into a conveniently brief survey. The methods adopted in the whole scheme of research are described and justified in what seems to be unnecessarily cumbersome and abstract language, though there is no doubt that good team-work co-ordinating several scientific disciplines was achieved.

D.W.H.

Theory and Problems of Social Psychology. By DAVID KRECH and RICHARD S. CRUTCHFIELD. London: McGraw-Hill. 1948. Pp. xvi+639. 27s. net.

Introduction to Social Psychology. By OTTO FRIEDMANN. London: Sylvan Press. 1950. Pp. 220. 10s. 6d. net.

Social Psychology. By ALFRED R. LINDESMITH and ANSELM L. STRAUSS. New York: The Dryden Press. 1949. Pp. xvi+549. \$4.50 net.

As the torrent of text-books on social psychology sweeps past it seems likely that *Theory and Problems of Social Psychology* will remain in view longer than most, and not merely on account of its bulk (much of which could well have been spared). It is significant as making a bold attempt to abstract from general psychology a few guiding principles that have direct applicability to an understanding of social life, and it therefore does something to bring together a disciplined psychological viewpoint and an interest in the problems of society. The authors put their chief faith in an analysis of the 'attitudes' and 'beliefs' expressed in social behaviour. They succeed in showing that much social activity can be analysed with the help of these concepts and made more intelligible through experimental work on the processes of acquiring and changing beliefs and attitudes. Good sections on attitude measurement, opinion polls and propaganda emerge naturally from this plan of treatment. A particularly useful discussion of suggestion offers a modern account of a social process that earlier workers emphasized without fully clarifying, and that has not received a close enough examination from recent social psychologists. The book suffers from the handicap of a poorly defined audience: it handles social psychology in a way original enough to deserve full presentation and critical attention at the technical level, but at the same time much of it (especially the later chapters on practical problems) is evidently meant for quite immature students who need a repetition of elementary and familiar material. In thus aiming simultaneously at two targets the book is, unfortunately in a numerous company at the present time.

Dr Friedmann's intended audience is evidently the adult education class, and some features of his book (including well-selected reading lists) will commend themselves to tutors. It begins with a simple discussion of general psychology, strongly influenced by Freudian views, and proceeds to many very short chapters on all the usual topics—co-operation, competition, leadership, class, the crowd, prejudice, etc.—giving all the usual opinions current among 'enlightened' people at the present day. Whether these opinions emerge compellingly from scientific facts and principles is the doubt confronting all such social psychology: one suspects that with a different political and cultural outlook rather different opinions could be equally well rationalized by something that looked equally like a social psychology.

Social Psychology, by Drs Lindesmith and Strauss, has a clearly defined audience of undergraduate students, and offers them a treatment of the subject that emphasizes scientific principles more than practical applications. Discussion of the latter is limited mainly to the two fields of sex relations and race relations, though there are briefer, rather anecdotal treatments of propaganda, rumour and crowd reactions under the heading of 'Social Change'. The merit of the book lies in its thorough account of the psychology of language as a basis for social psychology. This leads on to an account of social role-taking, especially as affected by language, and the importance of social roles to the definition of the self, a theme for which the work of G. H. Mead is mainly drawn upon. Chapters on personality in society, disorganization and conflict, and deviant behaviour are developed fairly successfully from the same standpoint. The book is written in a simple and popular style, but handles its topics seriously enough to make at least Parts I and II ('Language Behaviour' and 'Socialization') useful to university students.

Pastoral Psychology. A study in the Care of Souls. By GÖTE BERGSTEN. London: George Allen and Unwin. 1951. Pp. 227. 15s. net.

Pastoral Psychology, by Göte Bergsten, raises the sights so high that its failure to hit the target is not surprising. Nor can this be cause for blame; for the objective the book proposes, namely the establishment of a new science of pastoral psychology, is as yet only a remote possibility. Those who have written on the care of souls have shown no evidence that the problems in this field are being attacked with scientific methods other than the more familiar psychological techniques, or that any systematic effort is being made to study the ordinary phenomena of spiritual life in a scientific way. But Mr Bergsten is convinced that: 'A new science of pastoral psychology is appearing, developing as an independent, rational form of treatment for the whole personality, and claiming recognition alongside both psychological and physical medicine.' This betrays Mr Bergsten's inadequate conception of the requirements of science and psychology. Therein lies the reason for the unscientific and popular character of much of the book which will perhaps

irritate the professional psychologist, as will the neglect of some more recent psychological work on education, character and religious conversion. Here and there the argument is scrappy and the English uncertain.

These imperfections, however, are minor in comparison with the value of the total fund of sound knowledge and advice the book offers practitioners in the care of souls. It genuinely attempts to fill a serious gap in the equipment of the average priest or minister. Pastoral psychology is the cinderella of the theological colleges, often being taught by tutors with no specialized training in psychology. Bringing under levy a vast range of psychological, theological and devotional literature, Mr Bergsten's book offers something more than the popular review of psychological theory so common in books on pastoral psychology. It really provides clear and wise guidance upon the handling of almost every human situation the minister of religion is likely to meet.

Divided into six sections, entitled: 'The responsibilities and limitations of the task', 'Concerning confession', 'The psychology of unbelief', 'Guilt and fear of punishment', 'Symptomatic religiosity in neurosis and insanity', and 'The psychology of conversion and growth in grace', the work mentions almost every serious problem in the treatment of both normal and abnormal people. Particularly welcome is the excellent advice upon dealing with ordinary people and everyday matters. Mr Bergsten has avoided that preoccupation with the neurotic, characteristic of some popular books. Further, he has fully recognized the spiritual value of man and explored the possibilities of religious methods which go far beyond the limits of psychology.

W.L.J.

The Search for the Beloved: A Clinical Investigation of the Trauma of Birth and Pre-natal Conditioning. By NANDOR FODOR. New York: Hermitage Press. 1949. Pp. viii + 400. \$5.00.

The theme of this long and absorbing work, fully documented by dreams, has much in common with the Freudian view of the almost universal desire to get back to the womb. The author maintains that birth for every human being has a catastrophic reaction. This is shown either through the development of neurosis or dreams in which a longing is said to be clearly expressed for the pre-natal state of bliss. 'All Utopian visions reveal a nostalgia for the bliss which we lost on leaving the maternal womb.' He even goes so far as to maintain that 'Modern man satisfies his hunger for uterine bliss by flying in planes', as well as by climbing mountains, exploring the North Pole or diving into the sea.

It is obvious that sexual intercourse shows a desire in one sense to return both to the womb and the breast. But according to the author, whose vast experience I cannot equal, all neurosis without exception can be linked up with pre-natal existence and the subsequent trauma of birth. He does, however, admit that the life of the foetus is not one long joy-ride. There is an interesting chapter on the influence of maternal impressions on the developing infant, and even apart from the transmission of viruses through the blood stream, the author is inclined to believe in the possibility of other pre-natal influences. With regard to the foetal movements he quotes Dr J. Sadger as saying: 'There are in the foetus frequent and violent movements which have as their purpose the avoidance of painful feelings.' The mobility of the unborn might be just as well explained by the thesis that the infant is having a good stretch, or hitting back at the mother for some neglect on her part. Indeed, one could find support for this theory, as foetal movements are always much more vigorous when one is trying to go to sleep at night, proving that the child likes the rocking it gets when the mother is in motion!

There is little proof for the statements made throughout the book regarding the content of the foetal mind. He attributes to the infant a knowledge of the pubic arch during the process of birth!

While freely admitting that violent efforts at resuscitation of the new-born may have repercussions in later life, one must hesitate to accept the final implications of this imposing work which refers all human activity to the desire to get back to the womb. Moreover, it is difficult to believe that the new-born is experiencing vehement emotions when one sees the child with an expression of placidity on its face. There is no discussion of the effect of Caesarean section on the subsequent development of neurosis. Though interesting the book is over-long, and unfortunately has no index.

STELLA CHURCHILL

The Day Hospital: An Experiment in Social Psychiatry and Syntho-Analytic Psychotherapy. By JOSHUA BIERER. London: H. K. Lewis. 1951. Pp. 56. 6s. net.

'The Day Hospital does not belong to the era of Individual Psychiatry, nor to the era of Orthodox Psychoanalysis; it is part, or the forerunner, of an era which may be called tentatively the era of Social Psychiatry and Syntho-Analytic Psychotherapy.' Anyone who is inclined to take this seriously can go on

to read that the principles of Social Psychiatry and Syntho-Analytic Psychotherapy are: '(i) The Principle of Universality. (ii) The Principle of Relativity. (iii) The Principle of Conation. (iv) Centripetal-Centrifugal Interaction. (v) Multi-dimension Approach. (vi) Treatment through Experience.'

Coming nearer reality one reads that 'a Day Hospital may be defined as a place where an attempt is made to make available, as far as is possible, every type of treatment provided by a modern intra-mural department together with all the advantages and privileges enjoyed by patients in an extra-mural department'. The book is offered as a description of the work of such a day hospital that has been established since 1948, but very little precise information about the hospital is in fact given. We are not even told whether patients attend daily or for how many hours a day. A number of case histories are given. Many methods of treatment are used, with chief reliance apparently upon encouragement, exhortation and membership of therapeutic social clubs. Adler and Moreno seem to be the regnant influences. Although the report is so pretentious and unconvincing the clinic may, of course, be doing useful work.

The Psychology of Flight. By ALEX VARNEY. New York: D. Van Nostrand Company Inc. (London: Macmillan). 1950. Pp. vi+269. 28s. net.

'The most uncertain element in flying to-day is the pilot... we know what to expect of airplanes... But when we can predict equally well what pilots will do, aviation will be much safer.' This is the viewpoint of the author, a flying instructor of great experience. 'Our emotions', he writes, 'often betray us in a heart-breaking manner when the chips are down and the stakes are high.' His thesis is: 'We can prepare for effective action through a study of our nervous system and our emotions to learn how and why we respond as we do to the stimuli arising from our environment.' His advice is sound, e.g. 'When one knows what to do in an emergency when there is a clear plan for action in his mind, he [the pilot] does not succumb to fear'; or again, the flying instructor 'must plan the student's flying experience so that he will learn in easy, controlled steps... He must train the student to succeed—not fail... A student should never be placed in a situation that is impossible for him to master; he must be prepared for the difficult situation through mastering simpler situations that lead up to the difficult situation.' Psychologists will find much of interest in the common-sense discussions of the errors which pilots have made. It is to be hoped that the publishers in this country will find some way of providing a cheap edition, so that the book will be read widely by flying instructors and by those learning to fly.

D. RUSSELL DAVIS

Probability and the Weighing of Evidence. By I. J. GOOD. London: Charles Griffin. 1950. Pp. viii+122. 16s. net.

This is by no means an easy book. Most experimentalists who use statistics for practical purposes are likely to remain content to do so without examining the theoretical and mathematical presuppositions underlying their methods in such detail as Dr Good invites them to. Yet those more philosophically inclined may welcome this opportunity to clarify their minds on the theoretical aspects of the methods they use.

R.H.T.

Analytical Group Psychotherapy. By S. R. SLAVSON. Columbia University Press (London: Geoffrey Cumberlege). 1950. Pp. x+275. 22s. 6d. net.

In this book Mr Slavson describes the psychotherapeutic processes at work in groups composed of either children, adolescents or adults; the work described was done at the Jewish Board of Guardians Child Guidance organization in New York.

The first part of the study is devoted to a general theoretical discussion of the basic mechanisms of therapy and how they function when individuals are treated together. Mr Slavson shows clearly that a therapeutic 'group' is, in fact, a 'compresence of patients' rather than a group in the usual social sense. A detailed and interesting account of the actual group sessions follows. The role of the therapist is described with considerable acuity; both successes and failures are used as illustrations. The final chapters describe the possibilities of group psychotherapy in approved schools and mental hospitals.

The book as a whole is an informative account of a technique about which so far there is not a great deal published in this country. The sections on group work with children and in approved schools are especially interesting.

A.M.S.

Woman: Her Change of Life. By MIRIAM LINCOLN. London: Williams and Norgate. 1951. Pp. 116. 6s. net.

This readable little book should do much to break up the accretions of folk-lore and nonsense transmitted from the third and fourth generation to any particular present generation. Like the menstrual period, the menopause has been too often looked upon as a pathological, instead of a normal condition.

Dr Miriam Lincoln is a doctor with much experience and, in addition, an easy style of writing. She discusses the physiology of the life-changes, the symptoms that may or may not occur, and their psychological significance.

Throughout she insists that the menopause is a phase of development and not the end of all development, that life if properly lived can be very much worth the living, although the period of child-bearing is over. She does not, however, make the mistake of minimizing the discomfort that may, though not inevitably, occur. She points out further that changes of an analogous nature, though less obvious, occur in men, and she stresses the resulting need for mutual forbearance between husband and wife. M.S.

Les Variétés de la Constitution Physique de l'Homme. Introduction à la Psychologie Constitutionnelle. By W. H. SHELDON. Paris: Presses Universitaires de France. 1950. Pp. xii+382. 1000 fr. net.

Les Dimensions de la Personnalité. By H. J. EYSENCK. Paris: Presses Universitaires de France. 1950. Pp. vi+314. 700 fr. net.

These translations of books that are already well known in English have appeared in the series Bibliothèque Scientifique Internationale, which recently published a translation of Sir Godfrey Thompson's *Factorial Analysis of Human Ability* (listed in this *Journal* for March/May 1951).

OTHER PUBLICATIONS RECEIVED

- JALOTA, S. *Scientific Personnel Selection Procedure. A Study.* Hindu University, Benares. India: Dr S. Jalota. 1950. Pp. vi+80. Rs. 10s.
- JUNG, C. G. *Die Frau in Europa.* Zürich: Rascher Verlag. 1948. Pp. 39.
- JUNG, C. G. *Symbolik des Geistes.* Zürich: Rascher Verlag. 1948. Pp. xii+500.
- JUNG, C. G. *Die Beziehungen der Psychotherapie zur Seelsorge.* Zürich: Rascher Verlag. 1948. Pp. 39.
- KÜNZLI, ARNOLD. *Die Angst als abendländische Krankheit.* Zürich: Rascher Verlag. 1948. Pp. 292.
- LOEWY, HERTA. *The Retarded Child.* New enlarged edition. London: Staples Press. 1951. Pp. 160. 10s. 6d. net.
- NEUMANN, ERICH. *Tiefenpsychologie und neue Ethik.* Zürich: Rascher Verlag. 1949. Pp. 128.
- PEEL, E. A. *The Peel Group Test of Practical Ability.* Edinburgh: Thomas Nelson. 1951. Pp. 7. Single copy 7½d. net.
- THIEL, MANFRED. *Versuch einer Ontologie der Persönlichkeit.* Berlin: Springer-Verlag. 1950. Pp. xxiv+635. DM 29.70 net.